A Multi-objective Bi-level Optimisation model for Agricultural Policy in Scotland

Konstantinos Athanasiou Ververidis

A thesis submitted to the University of Edinburgh
for the degree of Doctor of Philosophy
Edinburgh 2006

Declaration

I hereby declare that the work presented in this thesis is the product of my own efforts, and has not been submitted in any previous application for a degree. The work on which it is based is my own except where stated in the text and in acknowledgement section.

Konstantinos Ververidis

To my Wife

Acknowledgements

I am sincerely grateful to my principal supervisor, Dr. Neil McRoberts for his invaluable help and support over the course of the research, especially at the "difficult" writing-up stage of the thesis. I also owe many thanks to Dr. Andrew Barnes for his supervision, support and friendship during the past three years. I am indebted to Prof. David Oglethorpe for believing in me and giving me the opportunity to work in the project.

I am deeply grateful to Dr. Deborah Roberts and Prof. Euan Phimister for their encouragement and understanding.

Also, many thanks go to all my friends in the Scottish Agricultural College in Edinburgh for reminding me that nothing compares to a friendly chat and a good laugh when things with the PhD go from bad to worse.

Finally I would like to thank my parents for all that they have done for me.

The research was funded by the Scottish Executive Rural Affairs Department.

Abstract

Agricultural policy analysis can be visualised as a multiple objective hierarchical optimisation problem whereby sequential non-cooperative interactions between the policy makers and the farmers take place. The objectives and choices of policy makers will almost always diverge from the objectives and choices of farmers. Policy makers exercise authority over some, but not all, of the variables in the total system whereas other variables affecting their multiple goals are under the direct control of myriad farmers who operate according to their own utility maximising motives. In order to advance their own objectives, the policy makers unilaterally and pre-emptively set the policy measures to influence the farmers. The farmers execute their decisions after, and in view of, the policies and make their production decisions that observe their goals best. Ultimately, the payoffs to both the policy makers and the farmers depend not only on the actions of the former, but also on the reactions of the latter. Such problems are difficult to solve due to their intrinsic nonconvexity and multiple objectives. This thesis shows how multi-objective genetic algorithms (MOGA) in conjunction with mathematical programming (MP) can be used for solving this type of problems. A MP model is developed to capture the production choices of farmers. The model is based on positive mathematical programming and its objective function parameters are estimated using the method of generalised maximum entropy. The model is nested in and controlled by a MOGA which captures the process of multi-objective optimisation of policy decisions. The approach is illustrated using a case study taken from the Scottish agricultural systems, where several socio-economic and environmental objectives for policy making are considered. Four types of policy instruments are examined: the current single payment scheme, a multi-payment scheme based on land use, an input taxation and a regulatory scheme. For a selection of scenarios alternative Pareto-optimal solutions are discovered and tradeoffs between the policy objectives are presented along with their associated production patterns. The performance of the modelling tool developed suggests that it is well suited to dealing with real-world policy issues. It offers considerable possibilities for exploring tradeoffs between non-commensurable and conflicting objectives relevant to sustainable development of Scottish agriculture.

Table of Contents

List of Abbreviations

Bi-level Programming (BLP)

Common Agricultural Policy (CAP)

Decision Maker (DM)

Elitist Non-Dominated Sorting Genetic Algorithm (NSGA-II)

Evolutionary Algorithm (EA)

Genetic Algorithms (GA)

Linear Programming (LP)

Mathematical Programming (MP)

Multi-objective Bi-level Programming (MOBLP)

Multi-objective Decision Making (MODM)

Multi-objective genetic algorithms (MOGA)

Multi-objective Optimisation Problem (MOOP)

Multi-objective Programming (MOP)

Positive Mathematical Programming (PMP)

Vector Evaluated *GA* (VEGA)

1. General Introduction

1.1. Introduction

This thesis is about policy optimisation. In fact, it is about agricultural policy optimisation towards multiple objectives. Optimisation is defined as an act, process, or methodology of making something (as a design, system, or decision) as fully perfect, functional, or effective as possible. The latter case is the view taken in this thesis. Like most of real life problems, multi-objective optimisation of agricultural policy too is a complex problem whose treatment requires the synthesis of different disciplines; elements of economics, environmental and agricultural science, as well as mathematical modelling are all involved in one way or another. The latter can be thought of as an interface device which makes those different disciplines work together in a compatible fashion. It also provides the building blocks for the construction of models of real systems which can be used to make inferences about how a real system might behave under hypothetical scenarios. The objective of this study is the development of a mathematical model of and for agricultural policy. With George Box's renowned quote "all models are wrong, some models are useful" in mind, the aim of the present undertaking is the development of modelling techniques and approaches that may be useful for agricultural policy planning and design. In particular the model should be able to increase the understanding of how agricultural policy interacts with agricultural systems and to explore possibilities for better policy design and implementation.

Agricultural policy and its optimisation is an interesting and practical subject worthy of investigating. Agricultural policy refers to a set of regulatory and economic instruments implemented in the agricultural sector. It aims to promote production patterns which ensure the sustainable development of rural economies and at the same time protect the agri-environment and maintain the natural resource base for future generations. Clearly, improving agricultural policy efficiency with respect to all these objectives offers significant benefits to society.

In terms of intellectual development, exploring ways to improve agricultural policy offers a unique challenge. Agricultural policy is a major driving force of the compound system economy-agriculture-environment. All the components are dynamically interlinked and interdependent. Changes in agricultural policy have an impact on the equilibrium state of the above triptych. For instance, different financial support regimes can produce different responses on behalf of the farmers reflected in their production choices and ultimately on the agri-environment. Simultaneously, agricultural policy is also shaped by these ever-changing conditions. The economy-agriculture-environment complex is, therefore, a typical example of a circular system of perpetual feedbacks where each component is an output and synchronously an input. Any effort to study the interrelationship between agricultural policy and economy-agriculture-environment triptych should take these facts under consideration. Overall, the study of agricultural policy optimisation is an intellectually stimulating pursuit with undoubted real life applicability.

1.2. The Problem of Agricultural Policy Analysis

The problem of policy design and analysis can be considered in terms of appraising alternative allocations of public resources toward multiple predefined objectives. The requirement that agricultural policy should integrate economic, social end environmental goals has been acknowledged in the latest reforms of the Common Agricultural Policy (CAP)[1]. In Europe therefore, designing the future agricultural policies should be: (1) cast in terms of "recognising" and promoting sustainable and multifunctional farming systems that optimise a suite of different and quite often conflicting socio-economic and environmental objectives and, (2) part of wider policies such as rural development.

These requirements pose a significant difficulty for policy makers and researchers since the level of achievement of the objectives is not dependent directly on policy regulations *per se*, but rather on the production choices of farmers arising under different market and policy conditions. Changes in production patterns have an effect

---

[1] Information and details on the latest and past CAP reform regulations can be found on the internet site of the European Union – EUROPA (http://europa.eu.int)

on a number of factors and functions of agricultural systems, and ultimately on the levels of achievement of policy goals. Policy makers have only a limited number of variables under their own control (taxes, subsidies, quantitative limitations, purchasing decisions etc.). Other variables affecting the government's goals and objectives are under the direct control of farmers who very probably have a different set of goals.

It is obvious that both the policy makers and the producers will adjust the variables that each of them controls in order to enhance their own utilities to the greatest extent possible. This distinction is important as when policy makers impose a particular policy, they may not have a precise idea of producer reactions. Furthermore, policy makers do not control producers' responses directly; rather they only guide them through the subsidy, taxation, and regulatory framework. It is important to maintain the distinction between government objectives and producer reactions. Within this framework, the efficiency (in terms of goal achievement) of any policy recommendation is subject to the response of the producers which is determined by their own objectives and goals.

Consequently, policy makers can not change production patterns directly, rather they have to achieve this indirectly by manipulating the options and changing the incentives available to farmers. The response of a sector to policy changes and the degree of adoption of the optimal production patterns that achieve the policy goals is mainly dependent on *i*) the responses of farmers to economic forces and signals (relative profitability of enterprises), and *ii*) personal objectives and preferences based on available resources and technology as well as environmental constraints. Hence, as O'Callaghan (1996) points out, "the individual objectives of farmers will only coincide with the aggregate criteria of the policy planner if the incentives are in place to bring about the required modification of production patterns".

1.3. Multi-objective Bi-level Optimisation of Agricultural Policy

What has been said so far suggests that the problem of policy optimisation has two distinctive attributes. The first attribute derives from the fact that agricultural policy is characterised by more than one objective or goal that the policy maker wishes to minimise/maximise or observe. These objectives refer to representations belonging to non-equivalent descriptive domains that cannot be reduced to each other. Also, different social actors have different and quite often conflicting values, perceptions and interests. This implies that a unique optimal solution cannot be defined, rather a set of alternative solutions, formally known as Pareto-optimal solutions, should be considered. In such cases, any decision is always associated with the generation of tradeoffs between the various objectives. To know the tradeoffs involved, the goal is to have available as many different Pareto-optimal solutions as possible. The existence of multiple objectives illustrates the multi-objective optimisation problem (MOOP) and algorithms for solving these problems should be able to find Pareto-optimal (or near Pareto-optimal) solutions.

The second attribute of the policy optimisation problem derives from the fact that there exist two interacting decision-making levels within a predominantly hierarchical decision-making structure. The execution of decisions is sequential from higher (the Leader) to lower level (the Follower). Each level possesses its own independent goals and objectives which may conflict with the goals of the unit at the other level. Each decision-making level exercises authority over some, but not all, of the variables in the total system. Neither level can directly change the variables under the control of the other level. The Leader unilaterally and pre-emptively sets the values of the variables under its own control to influence the Follower in order to advance its own objectives. The Follower executes its decisions after, and in view of, the policies of the Leader and sets its variables at the values that observe its goals best. Each level optimises its own objective functions independently of the other level, but affects and is affected by the actions and reactions of the other level. The external effect on a decision-making unit can be reflected in both its objective functions and its set of feasible decisions. This interplay between a Leader and a Follower effectively illustrates the bi-level optimisation or bi-level programming (BLP) problem, which is a special case of multi-level programming[2]. Therefore,

[2]A multi-level programming problem is a hierarchical optimisation problem with many decision-making levels, based on the less complicated Stackelberg game which assumes that decisions are

agricultural policy optimisation should be formulated as a multi-objective bi-level programming (MOBLP) model.

Following Hazel & Norton (1986), the problem of agricultural policy can be decomposed into the following two interacting problems:

(1) the multi-objective optimisation problem of allocation of public resources (policy optimisation), given predictions of how farmers will react to each possible allocation and the impact of their reaction on the achievement of the policy objectives;
(2) the problem of predicting farmers' responses to policy changes given their own objectives.

Suppose the set of relevant policy decision variables, $\mathbf{x}^{p}$ consists of economic instruments, $\mathbf{x}^{p1}$, and regulatory instruments, $\mathbf{x}^{p2}$. Also, the policy problem has N objective functions, $f^{p}_{n}(\mathbf{x}^{p}, \mathbf{x}^{f})$, and the farmers' problem has M objective functions, $f^{f}_{m}(\mathbf{x}^{p1}, \mathbf{x}^{f})$. All objectives are functions of either the policy or the farm decision variables and at least some of the objectives are functions of both types of decision variables. Policy objectives may include diverse outcomes such as reductions in soil loss, pesticides used, carbon emissions, government subsidy costs, and increases in farm employment and earnings by small farmers. Typical objectives of farmers are maximisation of profits and minimisation of risk. The problem of multi-objective bi-level agricultural policy optimisation may be expressed algebraically as follows:

Find the vectors of $\mathbf{x}^{p} = (\mathbf{x}^{p1}, \mathbf{x}^{p2}) \in \Re^{k}$ and $\mathbf{x}^{f} \in \Re^{i}$ that:

$$\max / \min \quad f^{p}_{n}(\mathbf{x}^{p}, \mathbf{x}^{f}), \qquad n = 1, 2, \ldots, N \qquad \text{(1a)}$$

Subject to

$$\mathbf{A}_{1}\mathbf{x}^{p} \leq \mathbf{b}^{p} \qquad \text{(1b)}$$

$$\mathbf{x}^{p(L)} \leq \mathbf{x}^{p} \leq \mathbf{x}^{p(U)} \qquad \text{(1c)}$$

$$\max / \min \quad f^{f}_{m}(\mathbf{x}^{p1}, \mathbf{x}^{f}), \qquad m = 1,2, \ldots, M \qquad \text{(2.a)}$$

---

made sequentially with no elements of bargaining, cooperation or effective bluffing (Martin *et al*, 2000)

Subject to

$$\mathbf{A}_2\mathbf{x}^{\mathrm{f}} \leq \mathbf{b}^{\mathrm{f}} \quad (2.b)$$

$$\mathbf{A}_3\mathbf{x}^{\mathrm{f}} \leq \mathbf{x}^{\mathrm{p}2} \quad (2.c)$$

$$\mathbf{x}^{\mathrm{p}}, \mathbf{x}^{\mathrm{f}} \geq 0 \quad (2.d)$$

Where $\mathbf{b}^{\mathrm{p}}$ is the vector of policy resources; $\mathbf{A}_1$ is the matrix of coefficients that express the policy resource requirement per unit of the policy variable; $\mathbf{x}^{\mathrm{p(L)}}$ and $\mathbf{x}^{\mathrm{p(U)}}$ are variable bounds restricting the values each decision variable x^{p}_k can take; $\mathbf{A}_2$ is the technological sub-matrix of per unit resource requirements not affected by policy constraints; $\mathbf{b}^{\mathrm{f}}$ is the vector of resource endowments and other constraints; $\mathbf{A}_3$ is the technological sub-matrix of unit resource requirements constrained by policy regulatory variables, $\mathbf{x}^{\mathrm{p}2}$.

Equations (1a)-(2d) represent the upper level policy optimisation problem; *i.e.* defining the policy measurements that best achieve a set of goals under constraints, given predictions of how farmers will react to those measurements. Equations (2a)-(2d) represent the lower (nested) farm optimisation problem; *i.e.* predicting how farmers' production activities are affected by changes in policy variables. It should be noted that decisions on any one activity are influenced by the effect of policy variables on all the activities because it is the relative profitability of production activities and the available resources that determine production choices.

Many published attempts to deal with some of the aspects of this bi-level problem use economic models (Jones *et al*, 1995; Moxey *et al*, 1995) or ecological-economic models (*e.g.* Oglethorpe and Sanderson, 1999; Topp and Mitchell, 2003) based on MP techniques. The main attraction of MP models for policy analysis is that they allow the specification of a wide range of price and non-price policy instruments (O'Callaghan 1996). Therefore, under various policy scenarios, the impact of policy changes on land use patterns, and effectively on the agri-environment, can be modelled. The policy makers can then use this information to compare the various alternatives and select the policy that best meets their objectives. The main disadvantage of such an *ad hoc* procedure is that even for a few scenarios and objectives the comparison becomes difficult and decision-making is subjective and uncertain. Hence, only a limited number of alternative policy designs can be examined leaving most of the space of feasible solutions unexplored (Rubenstein-Montano *et al*. 2000). Consequently, policy optimisation does not occur and better policy designs are likely to be missed.

In addition to standard MP models, various multi-objective programming (MOP) techniques have been developed and applied in order to support decision-making in agriculture (Romero and Rehman, 2003). Some studies (Zander and Kächele, 1999; Tiwari *et al*., 1999; Pacini *et al*., 2003) used MOP models to optimize the agri-environment; *i.e.*, to find out how farming systems should be configured in order to achieve multiple environmental-economic objectives. However, these studies assume that policy and producers' objectives coincide, and are subject also to the same limitations described before. In both cases, the policy problem reduces to one level and the policy solutions obtained are likely to be sub-optimal.

Such traditional modelling tools for economic policy analysis are each well suited to specific analytical tasks, though not specifically intended to incorporate the distinctly different perspectives, constraints and hierarchical interplay between decision makers at two different levels. Bi-level programming aims to fill this gap. There are a number of applications from a variety of fields that have been formulated as BLP problems. In an agricultural policy context, Candler and Norton (1977) and later Candler *et al* (1981) illustrate how bi-level programming can be used to analyse the dynamics of a regulated agricultural economy. Önal *et al* (1995) investigate improved allocation of subsidised credits among farm groups in Indonesia and, in a more recent study for the French Ministry of Agriculture, Bard (1996) used bi-level programming to examine the economics of promoting biofuel production from farm crops.

While BLP problems are obviously very common in economics and policy planning, bi-level modelling has not yet gained wide application. This is due, primarily, to the difficulty in solving realistic problems of this type because their geometric properties are more complex than standard mathematical programming. They are analytically difficult because they are combinatorial[3] in nature, have a generally nonlinear, non-convex feasible space(existence of many local optima), even when all functions are linear, so that traditional optimisation algorithms frequently fail to uncover the optimal solution (Martin *et al*, 2000; Colson *et al*, 2005). The major complexities in solving these problems are that the response function of the lower level to the upper-level decisions might not be defined uniquely and that, in general, the overall formulation is a nonconvex and nonsmooth mathematical programming model. In modelling and solving policy analysis problems, an additional complicating factor is the multi-objective nature of the policy problem.

No universal algorithm exists for the solution of bi-level programming problems (Candler, *et al*., 1981; Bard 1985; Bard 1998). A number of classical[4] algorithms have been proposed which can be divided into six different classes: Algorithms based on (1) *branch and bound* methods, (2) the *extreme point* based algorithms, (3) *complementary pivot* algorithms, (4) those using *descent* methods, (5) those using *penalty function* methods, and (6) those which are based on *trust-region* methods. The more interested reader is referred to Vicente and Calamai (1994) and Colson *et al* (2005) who provide introductory reviews of classical methods for BLP problems. There are also several studies dealing with the development of heuristic processes for solving bi-level and multi-level programming problems. Oduguwa and Roy (2002) and Liu (1998) propose genetic algorithms for the bi-level and the multi-level case respectively. Following the proliferation of studies devoted to bi-level programming, a number of textbooks have also been published. Among them, those by Bard (1998) and Dempe (2002) are maybe the most popular.

[3] Loosely speaking, combinatorial optimisation problem is any optimisation problem that has a finite number of discrete feasible solutions, a property which makes them harder to solve than problems with a continuous set of feasible solutions.
[4] Classical algorithms use the mathematical properties of the BLP problem to either reformulate it into simpler one or to assist the search by the algorithm

1.4. A simple approach for solving the MOBLP problem

Given the limitations associated with traditional algorithms for MOBLP and the lack of readily available solvers for BLP problems the only option, for this study, was to develop a method for solving the MOBLP policy model from scratch. Due to the complexity, effort and uncertainty involved with the development and implementation of a new mathematical algorithm, alternative approaches had to be considered. Therefore, the problem was not approached as a purely mathematical problem but rather in an “unconventional” way which is based on two observations. The first one is that policy optimisation can be thought of as a sequential process of continuous improvement of policy solutions similar to that of natural evolution. The second observation (which derives from the bi-level nature of the policy problem) is that this evolution process is subject to and depends on the farmers’ responses, which is in itself another optimisation procedure.

With these observations in mind, the solution to the MOBLP policy problem suggested in this study is an evolutionary-style optimisation for the agricultural policy problem with the optimisation of the farmers’ problem embedded[5].

By analogy with natural evolution, where species need to adapt to a complex and dynamic physical environment in order to persist, policies need to adapt to a complex and dynamic socio-economic environment. Different policy designs consist of different sets of policy variables in analogy with the way in which different genotypes consist of different sets of genes. In biology, the physical characteristics of the organism (*i.e.* the phenotype) determine how well suited to its environment a species or an individual organism is. This suitability is usually referred as “fitness”. In a policy problem context, fitness can refer to the level of achievement of the policy objectives, which, it is re-emphasised, is partly a result of the farmers’ production choices associated with each policy. In biological systems, fit individuals are also more likely to survive and reproduce passing on some of their “good” genes to their offspring. Also new genes may be created due to mutations (partly induced

[5] It was only at the final stages of this work when the author of this thesis became aware that a similar approach was already proposed by Yin (2002) for a MOBLP problem in transportation planning and management.

by the prevailing environmental conditions) during reproduction. Over the generations, species evolve; *i.e.* adapt, improve and become more efficient. Likewise, in any given set of different policy designs it is likely that some will be fitter than others. There is also a chance for combinations of the fittest designs to produce some new design(s) with even better fitness. If policy designs with better fitness are selected for combination and included in future generations of policy evolution more times than less fit designs, through repeated selection, the fitness of individual designs as well as the total fitness of the whole set is improved. After a sufficient number of combinations/generations the best policy designs can be found.

A similar pattern can be observed in the real evolution of agricultural policy preceding decades, where new measures seem to have emerged from old ones. For example, two of the most popular policy instruments of the CAP were the payment system based on output subsidies and the new agri-environmental measures which both belong to the “species” *direct payments.* However, only the latter, which has evolved from the former, seems well-suited to deliver the policy objectives that reflect the current socio-economic environment. Whether it will deliver the policy goals that it was originally aimed at depends a great deal on how the farmers will respond.

The analogy between optimisation and natural evolution is not new. The idea of mimicking the processes of natural selection and genetics to solve practical optimisation problems was considered several times during the 1950’s and 1960’s. These efforts, and particularly the work by Holland (1975), gave birth to a class of heuristic optimisation algorithms known as Genetic Algorithms. Genetic algorithms (GAs) are adaptive methods based on genetic and evolutionary processes of biological organisms used for solving search and optimisation problems. By mimicking the biological operations of natural evolution such as selection, reproduction and mutation, they capture mechanistically the main elements of this “optimisation” process.

Literature reveals that GAs are quite effective and robust in dealing with non-convex problems, and they do not require the objective function to be differentiable (Oduguwa and Roy, 2002). This is because they direct their search only based on the values of the objective functions and not on their mathematical properties. Hence, their objective functions need not be mathematical functions. Also, they can work in conjunction with other procedures such as computer programmes. This "hybridisation" ability of GAs makes them particularly suited to handling the task required in the study namely, embedding another optimisation algorithm in the GA. Genetic Algorithms have also been extended to handle optimisation problems with multiple objective functions. Since, for several simultaneous objectives, a single optimum solution (the so-called *ideal solution*) is not feasible, MOGAs seek to identify a set of *Pareto-optimal* solutions. Because no one solution is better than any other solution in the *Pareto-optimal* set (in the absence of any further information), it is also a goal to find as many such *Pareto-optimal* solutions as possible. MOGAs are more likely to identify multiple optimal solutions simultaneously because they work with populations of solutions than algorithms that seek for one solution (Deb, 2001). For all these reasons GAs can be applied successfully in almost any optimisation problem (Forrest, 1993).

Genetic algorithms have also been used as an alternative for solving bi- and multi-level programming problems. For example, Liu (1998) proposes a GA design for solving Stackelberg-Nash equilibrium of nonlinear multi-level programming with multiple followers in which there might be information exchange among the followers. More recently, Oduguwa and Roy (2002) developed a Bi-level GA (BiGA), which is a sequential nested algorithm. BiGa solves two optimisation problems iteratively, the first for the leader in all the x variables under his control and a subset of the y variables associated with the optimal basis follower's problem, and the second for the follower problem with all the x variable fixed. The follower problem is solved for y with x as parameters and the solution is returned to the leader problem. The algorithm uses the genetic operators for two populations, one for each problem. Haubelt *et al*, (2003) propose a novel multi-objective evolutionary algorithm for solving hierarchical multi-objective optimisation problems. To deal with the complexity of the search space they use a hierarchical structure for chromosomes. These are based on the idea of *regulatory genes* where the activation and deactivation of genes is used. New genetic operators are also defined to process the hierarchical chromosomes, namely *composite mutation* and *composite crossover*.

All the above approaches use the GA to solve the problems of all levels. Only Yin (2002), for a transportation planning problem, implements the idea of using a GA only for the upper level problem and a classical solver for the lower level problem. Apart from this difference the working principle of his method is similar to Oduguwa and Roy (2002). The approach by Yin (2002) appears simpler and likely to be more efficient because it (1) partitions the optimisation search into two separate optimisation stages one for each level and, (2) solves the lower level one with a classic algorithm such as the Frank-Wolfe algorithm. Yin (2002) found that the method was efficient to search simultaneously the Pareto-optimal solutions for the multi-objective bi-level transportation problem, providing (as he said) “a powerful, and possibly interactive, decision tool, allowing the decision-makers to learn about the problem before committing to a final decision”.

It follows that MOGAs are particularly suited for implementing the approach for solving the MOBLP policy problem examined here because they can be used for (1) performing the evolution-based multi-objective optimisation of the policy problem and, (2) embedding a second optimisation that solves the farmers’ problem separately.

The technique to implement the approach is conceptually quite simple. It borrows from two fundamental ideas in computer programming, namely (1) the modularisation of computer programmes and (2) their inter-operability. Such programmes utilising these ideas are divided in a number of modules each consisting of a block of programming code built to perform a specific task. The modules exhibit inter-operability, whereby the exchange of data and services between modules is possible. In the simplest case, each module can be called by any other module and executed; during this process, variables containing information are passed on from one module to the other(s).

The same principle can be applied to the BLP problem. The idea is to treat each optimisation as a different module that can communicate and exchange information with the other. Hence, two programmes are built; the first, called the follower module consists of the farm optimisation problem. It is responsible for three tasks. First, it receives values for the variables controlled by the second programme, called the leader module. Second, it uses these values to solve the inner optimisation problem by using a standard optimisation solver package. Third, it returns the solution to the leader module. The leader module consists of the policy optimisation problem. Apart from sending and receiving variable values it has also the more demanding task of searching for the optimal solution to the upper level optimisation problem. Because at every iteration of its search cycle the leader module exchanges information with an external process (i.e. the lower module and solver package), traditional mathematical algorithms are not suitable. Instead, a MOGA allows communication with other procedures at any stage of its search.

There are dozens of MOGAs freely and commercially available. Thus, there was not a need for developing one for the present project. The *Fast Elitist Non-dominated Sorting Genetic Algorithm for Multi-objective optimisation* (NSGA-II) developed by Deb *et al* (2002) was selected for the job. The choice was based on the following reasons: NSGA-II is fast, powerful, freely available, fairly user friendly and needed only a few modifications in order to be applied to the MOBLP problem treated in this study. Its full description is given in chapter 5.

For the farmers problem in the follower module a standard mathematical programming model was built. Its structure and the specification of its objective function are based on a calibration method called Positive Mathematical Programming (PMP) developed by Howitt (1995). PMP-based agricultural models calibrate to observations of a base year and have a smooth response to parameterisation. These merits make them very attractive for use in agricultural policy modelling. The PMP-based model is called SASEM (Scottish Agricultural Systems Economic Model). As the name suggests it is developed for the Scottish agricultural sector. It has a quadratic objective function with 88 decision variables and 45 linear constraints. Due to its simple structure it can be solved by standard quadratic programming solvers that use the *Primal* and *Dual Simplex* methods. This computational simplicity means that finding the optimum values for the decision variables of the follower problem in any individual happens almost instantaneously on a standard PC, before SASEM returns the values back to the upper module where the MOGA uses them to direct its search for the Pareto-optimal solutions to the multi-objective bi-level problem.

In summary, the approach for solving multi-objective and bi-level optimisation problems for agricultural policy proposed in the thesis is a model that integrates Genetic Algorithms and classical Mathematical Programming. The work by Yin (2002) provides some reassurance that "hybrid" algorithms like the one that is proposed here are promising tools for research involving multi-objective bi-level optimisation problems. Given that the purpose of the present model is **A**gricultural **POL**icy **O**ptimisation, the name APOLO was thought appropriate.

1.5. Objectives of the Research and Outline of the Thesis

In the first two of the preceding sections the scope and the aim of the research were defined and the main existing difficulties and limitations involved were highlighted. The need to overcome these problems was also stressed. The third section outlined the approach adopted in the study for tackling these problems. The central research objective of the study is to build a mathematical model of the system, Farming – Agri-environment – Policy, for Scotland. The model aims to provide rational and useful answers to research questions such as:

1. What are the possible effects to and responses by the agricultural sector in Scotland of proposed changes in policy measures?

2. What are the possible impacts of those responses on the agri-environment of Scotland?
3. Are there any feasible policies that are more efficient than those currently implemented and what are those?
4. Can we find the best possible solutions for agricultural policy in Scotland for any given set of policy objectives and farmers' objectives when those do not coincide?

In order to build a model that would be able to address these questions the following tasks need to be tackled:

1. the model should distinguish between policy makers' and farmers' objectives;
2. the model should partition its variables into those under the direct control of the policy maker and those under the direct control of the farmers;
3. the model should handle multiple objectives and find a diverse set of Pareto-optimal or near Pareto-optimal solutions;
4. the solutions should consist of values for all policy objective functions and all decision variables as well as of values for all farmers' objective function(s) and all decision variables;
5. the policy model should be flexible enough to incorporate a range of types of policy variables and its objective functions should be sufficiently realistic;
6. the farmers' model should be representative of the whole agricultural sector, *positive* rather than *normative* and reliable with respect to its simulations.

This dissertation consists of three main parts. The first part concerns the development of SASEM the Mathematical Programming model for the Follower Module of APOLO. The second part is an exposition of the main components of the Leader Module and its combination with the Follower Module. The final part includes the "real-world" application of the proposed modeling framework.

Part I consists of chapters two to four. Chapters 2 provides a background on Mathematical Programming problems with the emphasis on Linear Programming models and their limitations and difficulties as modelling tools for agricultural policy analysis. The issues of aggregation and calibration bias involved with regional LP models are pointed out. Chapter 3 takes these issues a bit further by presenting a literature review of a group of new methods, so-called Positive Mathematical Programming, that have been developed to overcome these limitations. Chapter 4 gives the description of the Mathematical Programming model of Scottish agricultural systems (SASEM) built for the Follower Module of the multi-objective bi-level programming model APOLO. A considerable emphasis is given to how a state-of-the-art Positive Mathematical Programming based method which uses Maximum Entropy econometrics has been applied for parameter estimation in SASEM's objective function. The chapter and part I conclude with the validation of SASEM where results on the model's goodness-of-fit to past observations are presented.

Part II commences with Chapter 5 which gives a theoretical background on multi-objective genetic algorithms. The basics of multiple objective programming are introduced and classical and evolutionary algorithms designed for multi-objective programming problems are presented. The focus of attention is on multi-objective genetic algorithms, especially on the one that the APOLO model uses in the Leader Module. Chapter 6 follows with the synthesis of SASEM and the multi-objective genetic algorithm into the multi-objective bi-level programming model APOLO, specifically designed to tackle the problem of agricultural policy optimisation. The remainder of the chapter makes a case on the validity of the model by presenting model's results in solving test problems used in the literature.

Part III includes the final three chapters (7-9) of the thesis. Chapter 7 presents an application of the policy optimisation problem through a case study for Scotland. The policy context is given first; then a description of appropriate policy designs based on suitable policy objectives and instruments tailored for the case study follows. The chapter concludes with the setting of specific policy scenarios aimed at the investigation of the research questions (1-5) posed in the study. Chapter 8 presents the findings and a sort discussion. Chapter 9 concludes the thesis with a general discussion of the main issues that have emerged throughout this research endeavour.

PART I

Multi-objective Bi-level Optimisation of Agricultural Policy:

The Follower's model

2. Linear Programming Models for Agricultural Policy

2.1. Introduction

The principal focus of agricultural policy analysis is to understand the decision-making process of producers and predict their response to policy changes. One of the primary analytical tools used toward this aim is modelling that involves the reconstruction of production supply functions. This can be done by means of mathematical programming (MP); a term which refers to a set of procedures dealing with activity analysis. A desirable feature of MP is the ability to capture both the agricultural characteristics of the system and the farmer's decision-making rationale (Hazell and Norton, 1986). In recent years, MP has evolved considerably, losing the features of a pure farm management tool. Presently, it is used as an instrument of policy analysis at the regional, national as well as EU level, with the objective of analysing the impact of agricultural policies on supply and on the socio-economic and environmental systems linked to the farming sector (Salvatici *et al*., 2001).

Linear Programming (LP) has proven to be a very powerful MP tool for building models that reflect producer constraints, opportunities, and objectives. LP models can be solved under varying assumptions about the policy environment that affects producers, *i.e.* "what if?" type of questions could be addressed. Some examples are: what would happen to the expected income and the management of agricultural land if (1) the suckler cow premium were to decrease? (2) the compensation for cereal production were to cease? These cases illustrate the value of LP in analysing agricultural policy measures at the farm level.

The purpose of this chapter is twofold. Firstly, it serves to provide a brief methodological background for the non-expert for comprehending the technical chapters that follow. Secondly, it aims to highlight the strengths and weaknesses of LP models of farming systems in order to point out the need for developing more appropriate MP models for policy analysis.

2.2. Linear Programming Models of Farm Systems

The general structure of LP models can be described as a system of equations including a linear objective function whose value is to be optimised (maximised or minimised) subject to one or more linear constraints. In agricultural models the objective function usually is the producer’s profit or production cost, while the constraints are concerned with the maximum availability of limiting factors and inputs of production such as land, labour, capital, policy and market constraints.

LP models are based on the theory of fixed output and coefficients, postulating that a finite number of goods can be produced, and also that the number of potential techniques to obtain a given good is also limited. Linear programming theory further assumes that, in contrast with reality, the entrepreneurs in the model have perfect knowledge of the relation between production factors, the amount of output obtained and the level of unit costs of each production factor and future market prices.

The primary aim of the LP model is to solve the farmer’s first problem: identifying the unknown quantity, x, of goods that maximise/minimise a profit/cost function. Algebraically, the standard form of a linear optimisation problem can be written as

$$\max Z = \sum_{i=1}^{I} g_i x_i \qquad (1a)$$

$$\text{s.t.} \quad \sum_{i=1}^{I} \alpha_{ij} x_i \leq b_j^{res} \qquad [\lambda^{res}] \quad \text{all } j = 1 \text{ to } J \qquad (1b)$$

$$x_i \geq 0, \qquad \text{all } i = 1 \text{ to } I \qquad (1c)$$

in which

Z	objective function
x_i	a set of the i model activities (decision variables)
i	index of model activities
g_i	gross margins of model activities
$\Sigma\alpha_{ij}$	is the matrix of technical coefficients else denoted **A**
b_j^{res}	available resources
j	index of model constraints
λ^{res}	shadow prices of resource constraints

The elements of technology matrix **A**, α_{ij}, are the technical coefficients of production, which reflect the proportional relation between production factors and goods obtained in each production process. They characterise the individual production processes and show the technology available to the farm. Each production process can be carried out up to a maximum level, and is associated with a gross return. The level of each process depends on its relative usage of limiting inputs and gross return. Each limiting input can be used in the process until it is exhausted, which constitutes the constraint on carrying out the process. The optimum solution of the model is a mixture of activities and their levels, which optimises the objective function and at the same time satisfies the set of constraints. The optimum solution is defined such that changing the set of activities by including or excluding one or more activities or by altering their levels will give a worse result in the objective function.

Linear programming is concerned with making the best choice among a set of decision variables given that the system is bounded by the constraints and driven by an objective. Figure 2.1 shows a conceptual system where the decision-maker (*e.g.* farmer) can choose between growing two crops (both measured in hectares). The farming system is bounded by the physical constraints of land and labour, which are in finite supply. Thus, any choice which can be made must not exceed the use of these resources (*i.e.* it must be below the line on the graph). The set of possible solutions is represented by the feasible solution space (shaded area in Figure 2.1).

The only remaining question to be answered is where best to be positioned within the feasible solution space. This is dictated by the objective of the enterprise and the relative coefficient value of each decision variable (activity) in terms of the objective. In the case of maximising profit, the optimal point would lie at the tangential point of an isoprofit line to the feasible solution space, which represents the farthest possible point from the origin (Figure 2.2). Any point on an isoprofit line will yield equal profit but have differing levels of decision variables. The gradient of the isoprofit line is given by the relative values (objective coefficients, g_i) of the decision variables. This logic is identical when more decision variables are added – each variable adds another dimension to the solution space hypervolume.

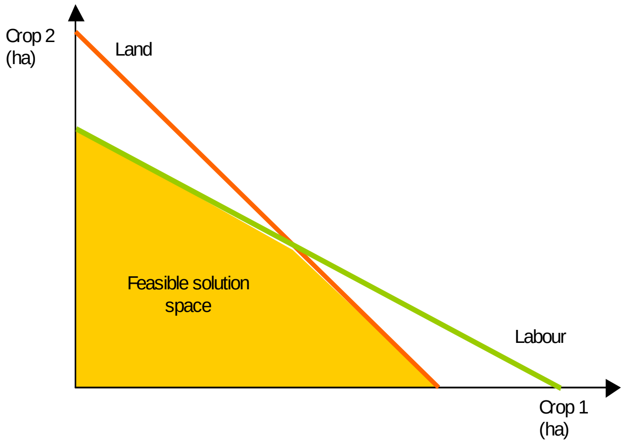

Figure 2.1. A two-dimensional representation of the LP rationale

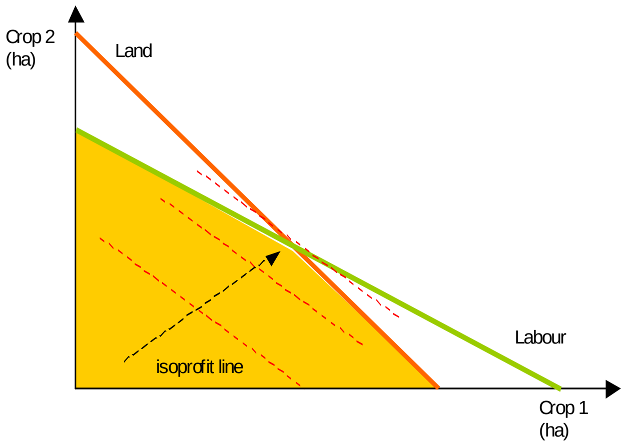

Figure 2.2. Conceptual model showing maximisation of profit

There is a symmetric problem in which we are able to analyse the 'hidden', less obvious, part of the farmers' decision-making problem. This is the problem inherent in the question: how much should the farmer be willing to pay to rent an additional unit of each resource? This is known as the dual problem. Once the quantities of factors available are known, the aim of the dual problem is to identify the cost of limiting resources in order to minimise the overall cost for the farm of maintaining economic equilibrium, postulating that unit cost is higher than or equal to the market price. The dual values, which are usually indicated by λ^{res}, express the unit cost of processes and offer indications on income changes if availability of the limiting actor increases or decreases by one unit. For this reason, the dual variables are also called shadow prices of the limiting inputs. It is possible to specify a dual linear programming problem to find the optimal shadow prices for the limiting resources.

2.2.1. Solving LP Models[1]

The matrix version of the basic LP problem can be expressed in the equations below.

$$\text{Max} \quad \Pi = \mathbf{g}'\mathbf{x} \tag{2a}$$

$$\text{s.t.} \quad \mathbf{Ax} \leq \mathbf{b} \tag{2b}$$

$$\mathbf{x} \geq \mathbf{0} \tag{2c}$$

where $\mathbf{x}$ is n × 1, $\mathbf{b}$ is m × 1 $\mathbf{A}$ is m × n.

An LP inequality system is commonly converted to equalities by adding so-called slack variables. These variables account for the difference between the resource endowment, $\mathbf{b}$, and the use of resources by the variables, $\mathbf{Ax}$ at no cost to the objective function. Thus,

$$\mathbf{S} = \mathbf{b} - \mathbf{Ax} \tag{3a}$$

$$S_i \geq 0 \quad \forall I \tag{3b}$$

Define the m × 1 vector of slack variables. Rewriting the constraints gives

$$\mathbf{Ax} + \mathbf{IS} = \mathbf{b} \tag{4}$$

where $\mathbf{I}$ is an m × m identity matrix. Slack variables appear in the objective function with zero coefficients. When the slack variables are added the resultant LP is

$$\text{Max} \quad \mathbf{g}'\mathbf{x} + \mathbf{OS} \tag{5a}$$

$$\text{s.t.} \quad \mathbf{Ax} + \mathbf{IS} = \mathbf{b} \tag{5b}$$

$$\mathbf{x}, \quad \mathbf{S} \geq \mathbf{0} \tag{5c}$$

Throughout the rest of the section the $\mathbf{x}$ vector is redefined to contain both the original x's and the slacks. Similarly, the new $\mathbf{g}'$ vector will contain the original $\mathbf{g}'$ along with the zeros for the slacks, and the new $\mathbf{A}$ matrix will contain the original $\mathbf{A}$ matrix along with the identity matrix for the slacks.

LP theory (Dantzig, 1963; Bazarra, *et al.* 1990) reveals that a solution to the LP problem will have a set of potentially nonzero variables equal in number to the number of constraints. Such a solution is called a *Basic Solution*, the associated

[1] The material presented in this section heavily draws on Howitt (2005b).

variables are commonly called *Basic Variables*, whereas the other variables, which are set to zero, are called the *Nonbasic Variables*. Once the basic variables have been chosen **A**, **g** and **x** can be partitioned as follows:

$$\mathbf{A} = \left[\mathbf{A}_B \vdots \mathbf{A}_{NB}\right], \mathbf{g} = \left[\mathbf{g}_B \vdots \mathbf{g}_{NB}\right] \mathbf{x} = \begin{bmatrix} \mathbf{x}_B \\ \cdots \\ \mathbf{x}_{NB} \end{bmatrix} \quad (6)$$

The system $\mathbf{Ax} = \mathbf{b}$ can be expressed as:

$$\left[\mathbf{A}_B \vdots \mathbf{A}_{NB}\right] \begin{bmatrix} \mathbf{x}_B \\ \cdots \\ \mathbf{x}_{NB} \end{bmatrix} = \mathbf{b} \quad (7)$$

Subsequently, the problem is partitioned to become

$$\text{Max} \quad \mathbf{g}'_B \mathbf{x}_B + \mathbf{g}'_{NB} \mathbf{x}_{NB} \quad (8a)$$

$$\text{s.t.} \quad \mathbf{A}_B \mathbf{x}_B + \mathbf{A}_{NB} \mathbf{x}_{NB} = \mathbf{b} \quad (8b)$$

$$\mathbf{x}_B, \quad \mathbf{x}_{NB} \geq \mathbf{0} \quad (8c)$$

Since $\mathbf{A}_B^{-1}$ exists by definition of a basis (m × m, m linearly independent rows), multiplying both sides of the constraints system by $\mathbf{A}_B^{-1}$ gives:

$$\mathbf{A}_B^{-1}\, \mathbf{A}_B\, \mathbf{x}_B + \mathbf{A}_\mathrm{B}^{-1}\, \mathbf{A}_{NB}\, \mathbf{x}_{NB} = \mathbf{A}_B^{-1}\, \mathbf{b} \quad (9a)$$

$$\mathbf{x}_B = \mathbf{A}_B^{-1}\, \mathbf{b} - \mathbf{A}_B^{-1}\, \mathbf{A}_{NB}\, \mathbf{x}_{NB} \quad (9b)$$

Equation (9b) is the first of the two fundamental equations of LP.

An optimal basic feasible solution (BFS) has the elements of $\mathbf{x}_B$ all non zero and all the elements of $\mathbf{x}_{NB}$ zero:

$$\mathbf{x}_B = \mathbf{A}_B^{-1}\, \mathbf{b} \quad (10)$$

The above equation shows how the basic solution variable values change. However, the point of optimising is to find, amongst all the basic feasible solutions, the one that optimises the value of the objective function. The *Simplex* solution approach relies on choosing an initial $\mathbf{A}_B$ matrix, and then interactively making improvements by changing the nonbasic variables from zero. Hence, the effect of those changes on the basis and on the value of the objective function is examined.

Accordingly, the partitioned **x** is used to write the partitioned objective function in terms of basis and nonbasis variables.

$$\Pi = \mathbf{g}_B' \, \mathbf{x}_B + \mathbf{g}_{NB}' \, \mathbf{x}_{NB} \tag{11}$$

substituting the r.h.s. of equation 10 into equation 11 for $\mathbf{x}_B$ yields:

$$\Pi = \mathbf{g}_B' \, (\mathbf{A}_B^{-1} \, \mathbf{b} - \mathbf{A}_B^{-1} \, \mathbf{A}_{NB} \, \mathbf{x}_{NB}) + \mathbf{g}_{NB}' \, \mathbf{x}_{NB} \tag{12}$$

factoring out $\mathbf{x}_{NB}$ gives the second fundamental equation of LP:

$$\Pi = \mathbf{g}_B' \, \mathbf{A}_B^{-1} \, \mathbf{b} + (\mathbf{g}_{NB}' - \mathbf{g}_B' \, \mathbf{A}_B^{-1} \, \mathbf{A}_{NB})\mathbf{x}_{NB} \tag{13}$$

($\mathbf{g}_B' \, \mathbf{A}_B^{-1} \, \mathbf{b}$) gives the BFS value of Π, $\mathbf{g}_{NB}'$ is the revenue contributed by a new activity from $\mathbf{x}_{NB}$ entering the basis, $\mathbf{g}_B' \, \mathbf{A}_B^{-1} \, \mathbf{A}_{NB}$ is the cost of moving the old basis values to fit in $\mathbf{x}_{NB}$ activities, and ($\mathbf{g}_{NB}' - \mathbf{g}_B' \, \mathbf{A}_B^{-1} \, \mathbf{A}_{NB}$) represents the net change in objective function value from moving new activities into the basis. In other words, ($\mathbf{g}_{NB}' - \mathbf{g}_B' \, \mathbf{A}_B^{-1} \, \mathbf{A}_{NB}$) is the (revenue – opportunity cost) of vector $\mathbf{x}_{NB}$, the incoming nonbasic vector. If one of the $\mathbf{x}_{NB}$ values is set to non zero, (i.e. it is brought into the basis), the objective function will be changed by this amount. Writing the second term of the equation in summation form yields

$$\Pi = \mathbf{g}_B' \, \mathbf{A}_B^{-1} \, \mathbf{b} + \sum_{i \in NB} (\mathbf{g}_i - \mathbf{g}_B' \mathbf{A}_B \alpha_i) x_i \tag{14}$$

If all but one (x_η) of the nonbasic variables are left equal to zero then this equation becomes:

$$\Pi = \mathbf{g}_B' \, \mathbf{A}_B^{-1} \, \mathbf{b} + (g_\eta - \mathbf{g}_B' \, \mathbf{A}_B \, \alpha_\eta) \, x_\eta \tag{15}$$

Since the first term on the r.h.s. of the equation is equal to the value of the current objective function, ($\overline{\Pi}$), then it can be rewritten as

$$\Pi = \overline{\Pi} + (g_\eta - \mathbf{g}_B' \, \mathbf{A}_B \, \alpha_\eta) \, x_\eta \tag{16}$$

For maximisation problems, the objective value will increase for any entering nonbasic variable if its term, $g_\eta - \mathbf{g}_B' \, \mathbf{A}_B \alpha_\eta$, is positive. The algorithm rule that is most commonly used to determine which variable to enter is: select the nonbasic variable that maximises the difference between revenue and opportunity cost per unit of the new activity $\boldsymbol{x}_\eta$. This is, the *Simplex Criterion Rule* of Linear Programming, and the term $g_i - \mathbf{g}_B' \, \mathbf{A}_B \, \alpha_\eta$ is called the reduced cost which is the net benefit of an activity η entering the basis. If there are no variables with positive reduced cost then the solution cannot be improved on and is optimal. Theoretically, the *Simplex* algorithm can fail to find the optimal solution if the LP is degenerate (*i.e.* when one of the basic variables in a basic feasible solution takes on a zero value).

Since the basis always has the number of variables equal to the number of constraints, then to put in a new variable one of the old basic variables must be removed. The *Minimum Ration Rule* of Linear Programming, explained below, dictates the variable to be removed from the basis when adding a new one. To examine what happens to the basis when the nonbasic variables are changed from zero, equation (9b) is used. Writing the second term of the equation in summation form yields:

$$\mathbf{x}_B = \mathbf{A}_B^{-1}\,\mathbf{b} - \sum_{i \in NB} \mathbf{A}_B^{-1}\,\alpha_i\,x_i \tag{17}$$

Equation (13) shows how the values of the basic variables are altered as the value of nonbasic variables change. Namely, if all but one (x_η) of the nonbasic variables are left equal to zero then the above equation becomes:

$$\mathbf{x}_B = \mathbf{A}_B^{-1}\,\mathbf{b} - \mathbf{A}_B^{-1}\,\alpha_\eta\,\boldsymbol{x}_\eta \tag{18}$$

This gives a simultaneous system of equations showing how all of the basic variables are affected by changes in the value of a nonbasic variable. Furthermore, since the basic variables must remain non-negative the solution must satisfy

$$\mathbf{x}_B = \mathbf{A}_B^{-1}\,\mathbf{b} - \mathbf{A}_B^{-1}\,\alpha_\eta\,\boldsymbol{x}_\eta \geq 0 \tag{19}$$

The nonbasic variable x_η may change until one of the basic variables becomes zero. Solving (2.5) for the first element of $\mathbf{x}_B$ to become zero, $\mathbf{x}_{Bi^*}$, gives

$$\mathbf{x}_{Bi^*} = (\mathbf{A}_B^{-1}\,\mathbf{b})_{i^*} - (\mathbf{A}_B^{-1}\,\alpha_\eta)_{i^*}\; x_\eta = 0 \tag{20}$$

and solving for x_η yields

$$x_\eta = \frac{(\mathbf{A}_B^{-1}\,\mathbf{b})_{i^*}}{(\mathbf{A}_B^{-1}\,\alpha_\eta)_{i^*}} \quad , \text{where } (\mathbf{A}_B^{-1}\,\alpha_\eta)_{i^*} \neq 0 \tag{21}$$

This shows the value of x_η which causes the i^* basic variable to become zero. To preserve non-negativity of all variables, the maximum possible value of x_η is

$$x_\eta = \frac{(\mathbf{A}_B^{-1}\,\mathbf{b})_{i}}{(\mathbf{A}_B^{-1}\,\alpha_\eta)_{i}} \quad , \text{where } (\mathbf{A}_B^{-1}\,\alpha_\eta)_{i} > 0 \qquad \forall\, i \tag{22}$$

The smaller the above ratio is, the higher is the allowable increase in x_η. Therefore, the activity which gives the minimum ratio is the one that is replaced in the basic solution.

2.3. Policy Modelling with Linear Programming

Presently, LP models are widely used instruments of policy assessment at the regional, national, as well as EU level, with the objective of analysing the impact of agricultural and economic policies on the socio-economic and ecological dimensions of the farming sector (Jones *et al*. 1995; Moxey *et al*. 1995; Hennessy, 2000). The main method for obtaining policy results from LP models is parametric analysis. This involves deriving empirical estimates of the *Supply Function* and the *Derived Demand Function* by changing the output prices or quantities of available inputs over a range of values. The two cases are presented separately below.

2.3.1. Generating Derived Demand Functions for Inputs

In order to generate the derived demand function for inputs the resource availability vector **b** is parameterised, that is, changed by small incremental values over a specified range. For each value the optimal solution is found and the shadow value of the resource, λ, is plotted against the amount of the input **b** to form the derived demand function.

Given the problem

$$\text{Max} \quad \Pi = \mathbf{g}'\mathbf{x} \tag{23a}$$

$$\text{s.t.} \quad \mathbf{Ax} \leq \mathbf{b} \qquad [\lambda] \tag{23b}$$

$$\mathbf{x} \geq \mathbf{0} \tag{23c}$$

the optimal solution is

$$\mathbf{x} = [\mathbf{x}_B \vdots \mathbf{0}]$$

$$\therefore \mathbf{x}_B = \mathbf{A}_B^{-1}\mathbf{b} \tag{24}$$

and a change in solution is given by

$$\Delta\mathbf{x}_B = \mathbf{A}_B^{-1}\,\Delta\mathbf{b} \tag{25}$$

If $\Delta\mathbf{x}_B$ is small enough so there is no change in the basis, the change in the objective function, $\Delta\Pi$, is

$$\begin{aligned}\Delta\Pi &= \mathbf{g}'_B\mathbf{x}_B - \mathbf{g}'_B(\mathbf{x}_B + \Delta\mathbf{x}_B) \\ &= \mathbf{g}'_B\mathbf{A}_B^{-1}\Delta\mathbf{b}\end{aligned} \tag{26}$$

redefining

$$g'_B\mathbf{A}_B^{-1} = \lambda$$

gives

$$\begin{aligned}\Delta\Pi &= \lambda\,\Delta\mathbf{b} \\ \therefore \frac{\Delta\Pi}{\Delta\mathbf{b}} &= \lambda \\ \therefore \frac{\Delta\Pi}{\Delta b_j} &= \lambda_j \quad \forall\ j\end{aligned} \tag{27}$$

Thus λ_j, which is called the shadow price, measures the marginal change in the objective function when availability of resource j changes, given no change in the basis A_B. But from the equation for λ ($\lambda' = g'_B A_B^{-1}$) it can be seen that the value of λ does not change with Δb unless (i) the basis A_B changes, or (ii) the objective function coefficients g_B are changed by parameterisation. This gives rise to the stepwise response to parameterisation shown in Figure 2.3 below.

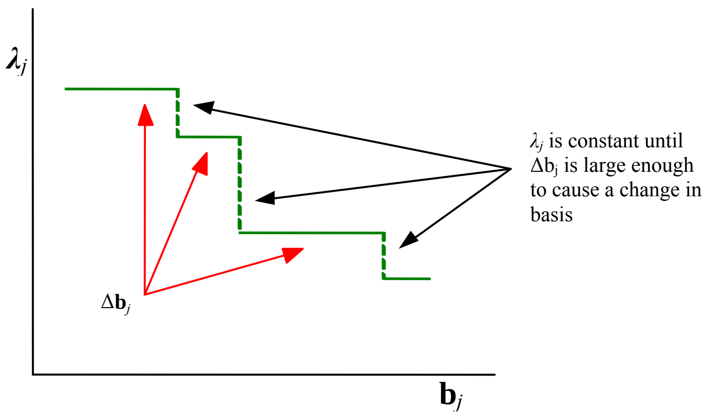

Figure 2.3. Step-wise derived demand functions for inputs

2.3.2. Generating Supply Functions of Outputs

Supply functions are obtained by parameterising the objective function coefficient, g_i, (*e.g.* price) of a single product by incremental values over a specified range. For each value the optimum solution to the model is found and the quantity of the output

x_i produced is plotted against the coefficient to form the supply function. Clearly, there will be a change in the product x_i when there is a change of basis caused by the change in g_i. The response of x_i when it is zero and therefore a non-basis activity is different from the situation when x_i is positive and is in the basis. Thus, there are two different cases:

Case I – Non basis activity supply parameterisation

According to the optimality conditions the reduced cost, g_i - $\mathbf{g}_B' \mathbf{A}_B \alpha_i$, for a non-basis activity, x_k, is negative. If g_i is increased so $g_i + \Delta g_i > \mathbf{g}_B'\mathbf{A}_B\alpha_i$ the reduced cost becomes positive and this will induce a change of basis which will bring the x_k activity into the basis with positive value as shown in Figure 2.4.

Case II – Basis activity supply parameterisation

For a basis activity, x_i, it is known that the reduced cost, g_i - $\mathbf{g}_B'\mathbf{A}_B\alpha_i$, is zero. Higher g_i value in the basis value vector will increase the reduced cost and the basis will have a larger value for x_i.

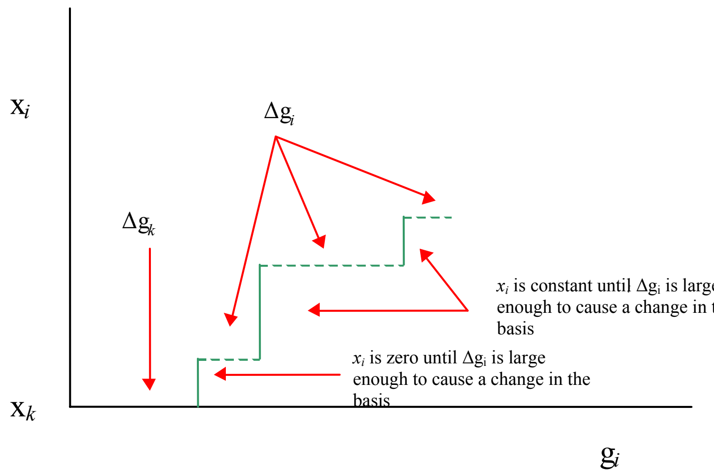

Figure 2.4. Step-wise supply functions of outputs

2.4. Limitations and Problems of LP Models

When building regional or sector models, ideally, a model for every individual farm should be constructed, then linked together to form the aggregate model. However, in practice this is not feasible and the regional model must be based either on models of representative farm types or on single macro-farm models. In the transition from small and medium scale to sector and national level analysis, when different production lines and regions are represented by 'representative farms' of homogeneous groups into which the farms are classified, aggregation bias (see Chapter 4: 4.2) arises unless, according to Day (1963), the farms satisfy the following three conditions:

- Technological homogeneity (each farm has the same production possibilities, the same type of resources and constraints, the same levels of technology, and the same level of managerial ability),
- "pecunious" proportionality (individual farmers hold expectations about unit activity returns that are proportional to average expectations), and
- "institutional" proportionality (the constraints for each individual farm should be proportional to the constraints of the aggregate farm).

The aggregate regional approach involves aggregating the resources of a homogenous region or area (not necessarily contiguous land) and modelling them as a single macro-farm. This approach carries the implicit assumption that all the aggregated farms have the same technology of production. In this case Hazel and Norton (1986) show that aggregation bias is always in an upward direction, namely, resource mobility is overstated by enabling farms to combine resources in proportions not available to them individually. One other common form of aggregation bias, overspecialisation, is said to exist if at the aggregated regional or sector level, observed production appears to be more diversified than in the outcomes of sector-level LP models. This happens because given the natural and economic conditions individual farms may specialise in production according with their resource constraints and preferences.

Overspecialisation prohibits LP models from calibrating to observed situations. The inability of LP models to be calibrated reduces their appeal as modelling aids for prediction purposes. Heckelei (1997) recognises three main reasons why the activities in the optimal solution of regional LP models are less than the number observed empirically:

- The number of empirically justified constraints relative to the number of observed production activities is smaller compared with the farm level
- Data, time and computational restrictions often do not allow to appropriate specification of the whole set of alternative linear activities with stepwise constraints, which would force the model to reproduce the base year solution.
- Nonlinear terms reflecting price formation or risk specifications, which would generally improve the solution, are often not incorporated into the model's objective function.

In calibrating LP models some parameters are changed in such a way that the model outcomes are close to actual data in terms of production quantities. One may, for example, add some linear or nonlinear terms to the cost function, or add risk aversion parameters to the objective function (*e.g.* Hazell, 1982). However, these approaches cannot guarantee exact model calibration. Another way to deal with overspecialisation is to extend the set of constraints, (crop rotation constraints, flexibility constraints, calibration constraints) until the total number of constraints at least equals the number of base period activities actually observed. Although that guarantees exact calibration it makes LP models inflexible because they remain subject to the same calibration constraints even under different scenarios runs where the initial baseline conditions are no longer prevailing.

2.5. Discussion

Critics of LP models believe that they are not very effective predictive tools due to their normative nature. However, Jones (1982), among others, argues that normative models have particular value in projecting response under hypothetical conditions outside the range of past experience. Such models indicate the likely direction of changes in production in response to price and policy changes. Furthermore, policy instruments such as dairy quotas, set-aside obligations in conjunction with crop specific payments per ha, stocking density restrictions, farm specific ceilings and the nitrate directive are much easier to model in a linear programming framework than in some others (Hennessy, 2000).

LP models are deterministic and assume instantaneous response to new conditions. As certainty and instantaneous response may not actually exist the optimal outcome may not mirror the actual. In addition, the linearity in programming models results in the following empirical problems. First, the methods used to calibrate linear models against the base-year data have to strike a balance between poor base-year calibration and fully constrained models that may bias policy results. Second, changes in the input prices or commodity support prices in the model do not cause changes in the shadow prices of inputs or output patterns unless they precipitate a change of "basis" solution. This leads to the stepwise response of LP models to parameterisation (Figures 2.3, 2.4). For models based on aggregate data, the range between steps may be larger than many levels of policy change, thus making the models inflexible for some types of policy analysis.

The most delicate aspect and biggest limitation of LP models is obtaining good estimates of the parameters that describe the technology used by and the objectives of farmers in order to adequately reproduce their aggregate decision-making process, given limitations on data availability. This chapter concludes by emphasising the need for developing ways to overcome the drawbacks of LP models. The next chapter offers a thorough review of methods aiming at calibrating MP models.

3. Calibration and Parameter Estimation of MP Models

3.1. Introduction

As noted in Chapter 2, LP models are optimisation models in which both the objective function and the constraints are linear. We saw that they make a useful tool for farm level decision making when relevant data are available. However, when data are limited, for reasons explained in Chapter 2 they do not calibrate to observed data. As Howitt (1995) shows, if a LP model does not calibrate to observed production activities with the full set of general linear constraints that can be empirically justified, a necessary condition is that the objective function be nonlinear in at least some of the activities. Following this 'nonlinear calibration proposition' he suggests an alternative calibration method called Positive Mathematical Programming (PMP) as a remedy for some of the problems associated with the LP approach. PMP is a 3-step process aiming to specify a nonlinear objective function whose optimal solution calibrates to a baseline state.

Among the advantages associated with the PMP approach are an exact representation of the reference situation, lower data requirements, and continuous (smooth) response to changes in exogenous variables.

The literature clearly suggests that the PMP approach is being adopted quite rapidly for agricultural regional and sector models (Heckelei et al, 2000). A comparative study by Arriaza and Gómez-Limón (2003) on the performance of widely used MP models reveal that models employing the PMP technique outperform classical LP and quadratic programming models. Nevertheless the technique itself has several limitations (Heckelei, 1997) the major one being the arbitrary specification of some of the objective function's parameters. Recent extensions of the method relax some of these limitations (Paris and Howitt, 1998; Heckelei and Britz, 2000; Paris, 2001;

Röhm and Dabbert, 2003). These mainly involve the development of econometric procedures for the estimation of the model's parameters.

The main objective of this theoretical chapter is to review and evaluate PMP and its extensions as well as related methods aiming at the specification, calibration and estimation of constrained optimisation models.

3.2. Calibrating with PMP

Positive mathematical programming is a methodology developed to calibrate linear programming models almost exactly to observed levels of agricultural activities in terms of output, input use, objective function values and dual values by transforming the linear objective function into a nonlinear one. The basic idea behind PMP is that it is easier to collect information on activity output of a farm than on production costs, and it therefore uses this information to construct models that can correctly represent the entrepreneur's observed behaviour.

The method assumes that some production activities are characterised by decreasing marginal returns. In order to derive rules for quantitatively specifying an objective function that expresses decreasing marginal returns, this assumption needs explanation. It refers to the observation that not only one output, (*i.e.* the crop associated with the highest gross margin as specified in an LP model), but a variety of outputs, are usually produced. Howitt (1995) argues that if rational behaviour is assumed for the entrepreneurs, then there must be an economic reason to undertake each production activity only to a certain extent before undertaking a different one.

According to the PMP rationale, the rational basis associated with the observed variety of production activities lies with the hidden cost associated with the production of each output. Hidden costs, causing a decrease in marginal gross margins, are viewed as a consequence of any factor that might contribute to increasing costs in general. Hidden costs refer to costs that are only perceived by the producers, while from a modeller's point of view they cannot be observed directly.

However, they can be seen indirectly from, for example, a cropping pattern, as it is assumed farmers are aware of the full amount of production costs and only grow a crop as long as it remains more profitable than other crops.

3.2.1. Standard PMP Approach

The standard PMP calibration approach, which can be attributed to Howitt (1995), uses three stages. These are described below.

3.2.1.1. Stage 1:Estimation of λ^{rec}, and λ^{cal}

The aim of PMP is to specify a nonlinear objective function which at optimum calibrates to a base year. In order for the parameters of this function to be specified information on the shadow values of calibrated activities is needed. The first stage provides this information.

A conventional linear optimisation model is extended by a set of calibration constraints for the given base year production level $\mathbf{x}^R$ and reformulated as follows.

<u>LP model 1</u>:

$$\max \quad f(\mathbf{x}) = \mathbf{g}'\mathbf{x} \tag{1a}$$

$$\text{s.t.} \quad \mathbf{Ax} \leq \mathbf{b}^{res} \quad [\lambda^{res}] \tag{1b}$$

$$\mathbf{Ix} \leq \mathbf{x}^R + \boldsymbol{\varepsilon} \quad [\lambda^{cal}] \tag{1c}$$

$$\mathbf{x} \geq 0, \tag{1d}$$

in which

- $\mathbf{g}$ is the vector of gross margins
- $\mathbf{I}$ is the identity matrix
- $\mathbf{b}^{res}$ is the vector of resource constraints
- $\mathbf{x}^R$ is the (k x 1) vector of the observed values of $\mathbf{x}$
- λ^{res} the vector of shadow prices of resource constraints
- λ^{cal} the vector of shadow prices of calibration constraints
- ε is a vector of small positive numbers added to the calibration constraints to prevent degeneracy.

The dimension of $\mathbf{x}$ is n x 1 and that of $\mathbf{A}$ is m x n, (m < n) and the basis dimension of $\mathbf{A}$ is m x m. At the observed optimal solution (nondegenerate in primal and dual specifications) there are k (k < n) non-zero values of $\mathbf{x}$. Therefore, in addition to the $\mathbf{x}_m$ basis activities of the optimal solution of the LP problem, there are an additional p activities $\mathbf{x}_p$, where p = k – m, that are observed and need to be calibrated into the optimal model solution.

The model has a very simple structure, as there are none of the constraints present in 'classical' LP models such as the constraints of crop rotation, output, sales etc. Instead it has two types of constraints: "structural" and "calibration". The first is to respect the overall availability of the land factor for the farm, and the second is to respect the production choices made by the farmer in terms of output quantities. Each restriction is associated with a corresponding shadow price (or dual variable): The structural constraints are associated with the shadow price vector $\boldsymbol{\lambda}^{rec}$, and calibration constraints are associated with the vector of differential marginal costs $\boldsymbol{\lambda}^{cal}$.

These values are estimated following the procedure developed by Howitt (1995), from the above auxiliary linear program (Problem 1). In this program the $\mathbf{A}$ matrix is partitioned into an m x m basis matrix $\mathbf{A}^m$ that corresponds to the m least profitable "marginal" activities ($\mathbf{x}_m$), which are upper bounded by the resource constraints, and an associated m x p matrix $\mathbf{A}^p$ for the p calibrated "preferable" activities ($\mathbf{x}_p$). Dropping out the n – k zero activities the optimal basic solution to the calibrated problem can be written as:

LP model 2

$$\text{max.} \quad \mathbf{g}_m'\mathbf{x}_m + \mathbf{g}_p'\mathbf{x}_p \tag{2a}$$

$$\text{s.t.} \quad \mathbf{Ax} = \mathbf{b} \qquad [\boldsymbol{\lambda}^{res(LP)}] \tag{2b}$$

$$\mathbf{x} = \mathbf{x}^R + \boldsymbol{\varepsilon} \qquad \left[\lambda_p^{cal} : \lambda_m^{cal}\right] \tag{2c}$$

or partitioned as

$$\begin{bmatrix} \mathbf{A}^m & \mathbf{A}^p \\ \mathbf{0} & \mathbf{I} \end{bmatrix} \begin{bmatrix} \mathbf{x}_m \\ \mathbf{x}_p \end{bmatrix} = \begin{bmatrix} \mathbf{b}^{res} \\ \mathbf{x}^R + \varepsilon \end{bmatrix} \tag{3}$$

The Lagrangian for the constrained LP problem 2 can be written as:

$$L = \mathbf{g}_m'\mathbf{x}_m + \mathbf{g}_p'\mathbf{x}_p + \boldsymbol{\lambda}^{res(LP)\prime}(\mathbf{b}^{res} - \mathbf{A}^p\mathbf{x}_p - \mathbf{A}^m\mathbf{x}_m) + \boldsymbol{\lambda}^{cal}((\mathbf{x}^R + \varepsilon) - \mathbf{x}) \tag{4}$$

The appropriate first-order optimality conditions are:

$$\frac{\partial L}{\partial \mathbf{x}^{p}} = \mathbf{g}_{\mathrm{p}} - \boldsymbol{\lambda}^{\mathrm{res(LP)}\prime}\mathbf{A}^{\mathrm{p}} - \boldsymbol{\lambda}_{\mathrm{p}}^{\mathrm{cal}} = 0 \quad (5a)$$

$$\frac{\partial L}{\partial \mathbf{x}^{m}} = \mathbf{g}_{\mathrm{m}}' - \boldsymbol{\lambda}^{\mathrm{res(LP)}\prime}\mathbf{A}^{\mathrm{m}} = 0 \quad (5b)$$

therefore solving for $\boldsymbol{\lambda}^{\mathrm{cal}}$ and $\boldsymbol{\lambda}^{\mathrm{res(LP)}}$ respectively yields:

$$\boldsymbol{\lambda}^{\mathrm{cal}} = \mathbf{g}_{\mathrm{p}}' - \mathbf{A}^{\mathrm{p}\prime}\boldsymbol{\lambda}^{\mathrm{res(LP)}} \quad (6a)$$

$$\boldsymbol{\lambda}^{\mathrm{res(LP)}} = (\mathbf{A}^{\mathrm{m}\prime})^{-1}\,\mathbf{g}_{\mathrm{m}} \quad (6b)$$

The right hand side of equation (6b) is the net change in the objective function value from moving into the expanded *Basis*, $\mathbf{A}^{\mathrm{m}}$, a less profitable 'marginal' activity, x_{m}. In other words, $\boldsymbol{\lambda}^{\mathrm{cal}}$, is the vector of opportunity costs of restricting the calibrated activities, $\mathbf{x}_{p}$, in the expanded *Basis*.

If a nonlinear function of an activity that needs to be calibrated, x_{p}, (which decreases the objective function value with the level of activity x_{p}, *e.g.* a nonlinear cost function) is added to the objective function, the marginal and average cost of producing x_{p} will differ. The net returns to land from x_{p} decrease as its level increases until they reach an internal equilibrium solution at the point where they are equal to the microeconomic *"equimarginal"* principle of optimal input allocation across activities. If the calibration constraints are removed and a nonlinear, say, (without loss of generality) cost function is added, LP model 2 becomes:

$$\max. \quad g(\mathbf{x}) = \mathbf{g}_{\mathrm{m}}'\mathbf{x}_{\mathrm{m}} + \mathbf{g}_{\mathrm{p}}'\mathbf{x}_{\mathrm{p}} - \boldsymbol{C}_{\mathrm{p}}(\mathbf{x}_{p}) \quad (7a)$$

$$\text{s.t.} \quad \left[\mathbf{A}_{B} \vdots \mathbf{A}_{NB}\right]\begin{bmatrix}\mathbf{x}_{B}\\ \cdots \\ \mathbf{x}_{NB}\end{bmatrix} = \mathbf{b} \quad (7b)$$

The specification of this nonlinear objective function is the concern of the second stage of the method.

3.2.1.2. Stage 2: Specification of the nonlinear function

The objective at the second stage of PMP is to specify the nonlinear objective function of the optimisation model (7a)-(7b). Shadow prices of calibration constraints (λ^{cal}) are used to derive the nonlinear parameters, which enter into the objective function. For the LP model 1 the Kuhn-Tucker conditions[1] of optimality at the base year situation yield:

$$\nabla f(\mathbf{x}^R) - \mathbf{A}'\lambda^{res} - \lambda^{cal} = 0 \qquad (8)$$

where $\nabla f(x)$ is the gradient of *f(x)*.

For the unconstrained (λ^{cal} no longer defined) PMP model with the nonlinear objective function, *g(x)*, the condition for the optimum at the same base year is

$$\nabla g(\mathbf{x}_R) - \mathbf{A}'\lambda^{res} = 0 \qquad (9)$$

Because optimum conditions of both models (calibrated LP/unconstrained PMP) refer to the same base year, which is exactly the idea when calibrating the PMP model to the LP model, the following equation also holds:

$$\nabla f(\mathbf{x}^R) - \mathbf{A}'\lambda^{res} - \lambda^{cal} = \nabla g(\mathbf{x}^R) - \mathbf{A}'\lambda^{res}$$

and if equal λ^{res} values are assumed for the two models then the above equation becomes

$$\nabla f(\mathbf{x}^R) - \lambda^{cal} = \nabla g(\mathbf{x}^R) \qquad (10)$$

In principle, any type of nonlinear function with the required curvature properties (convexity) qualifies for this step. For reasons of computational simplicity and lacking strong arguments for other type of functions (Howitt, 1995), the standard PMP method usually employs two types of quadratic specifications: a quadratic yield function and a quadratic cost function.

3.2.1.2.1. First Specification

A probable source of nonlinearity is the heterogeneous land quality, which leads to declining marginal yields as the proportion of a crop in a specific area is increased. This phenomenon, which was first formalized by David Ricardo, is widely noted by farmers but often omitted from quantitative production models.

Defining the nonlinear yield function as

$$y_i = \beta_i - \frac{1}{2}\delta_i x_i \qquad (11)$$

[1]A system of equations and inequalities which the solution of a mathematical programming problem must satisfy when the objective function and the constraint functions are differentiable.

then the quadratic total output function g(x) reads

$$g(x) = \sum_{i=1}^{p}(P_i(\beta_i - \frac{1}{2}\delta_i x_i)x_i - c_i)x_i$$ [2]

where P_i is the price of activity i's output and c_i is its variable cost.

Two equations are solved for the two unknown vectors of yield parameters the $\boldsymbol{\beta}$ and $\boldsymbol{\delta}$. The first equation is the vector of average yields for crops, $\bar{y}$

$$\bar{y} = \boldsymbol{\beta} - \boldsymbol{\delta}\mathbf{x}^R$$

The second equation is the equation (10)

$$\nabla f(\mathbf{x}^R) - \lambda^{cal} = \nabla g(\mathbf{x}^R)$$

which after substituting for $f(\cdot)$ and $g(\cdot)$ yields:

$$P\boldsymbol{\beta} - \boldsymbol{c} - \lambda^{\mathrm{cal}} = P\boldsymbol{\beta} - P\boldsymbol{\delta}\mathbf{x}_R - \boldsymbol{c}$$

Solving for $\boldsymbol{\delta}$ reads:

$$\boldsymbol{\delta} = \frac{\lambda^{cal}}{P\mathbf{x}^R} \tag{12}$$

When data for average crop yields, output prices and variable costs are available equations (11-12) is a system of two simultaneous equations with two unknowns; the vectors of parameters $\boldsymbol{\beta}$ and $\boldsymbol{\delta}$.

3.2.1.2.2. Second Specification

Defining yields per hectare as constant and marginal costs per hectare as increasing, such as the total variable cost function be a quadratic function, yields:

$$g(x) = \sum_{i=1}^{p}(P_i\bar{y}_i x_i - (\alpha_i + \frac{1}{2}\gamma_i x_i)x_i) \tag{13}$$ [3]

where P_i is the price of activity i and $\bar{y}$ is its average yield.

In the same manner the vector of the parameters of the cost function $\boldsymbol{\alpha}$ and $\boldsymbol{\gamma}$ can be derived by equation (10) and the following equation

$$\boldsymbol{P}\bar{y} - \boldsymbol{c} = \boldsymbol{P}\bar{y} - (\boldsymbol{\alpha} + \frac{1}{2}\boldsymbol{\gamma}\mathbf{x}^R), \tag{14}$$

---

[2] The linear objective function in this case is $f(x) = \sum_{i=1}^{p}(P_i\beta_i - c_i)x_i$, where yield $y = \beta$.

[3] The linear objective function in this case is $f(x) = \sum_{i=1}^{p}(P_i\bar{y}_i x_i - c_i x_i)$

which states that average gross margins are the same in both models. Substituting for *f(·)* and *g(·)* in equation (10) yields

$$\mathbf{P\bar{y}} - \mathbf{c} - \lambda^{cal} = \mathbf{P\bar{y}} - (\boldsymbol{\alpha} + \boldsymbol{\gamma}\mathbf{x}^{R})$$

which after rearranging becomes

$$\because \gamma = \frac{2\lambda^{cal}}{\mathbf{x}^{R}} \qquad (15)$$

Substituting $\boldsymbol{\gamma}$ in equation (14) yields the solution for $\boldsymbol{\alpha}$.

With both specifications the PMP models calibrate exactly to the observed activity levels. An empirical check on the calibration is performed by calculating the value marginal product (VMP) of the activities (Howitt, 1995). If they are close and converging, the PMP will calibrate without the additional calibration constraints.

3.2.1.3. Stage 3: Calibration to a base year

In this last stage, the calibration constraints of the first stage are removed and the model with the nonlinear objective function as specified in stage two is solved and typically calibrates almost exactly with the base year activity levels. The example that follows demonstrates the PMP calibration procedure.

3.2.1.4. Numerical Example of applying PMP

Let be $x^R = (x_w, x_o)$ the vector of the areas where two crops, wheat and oats, are grown on a farm in Scotland, and the gross returns for each of these crops are:

$c_w = 2.98 \text{ x } 69 - 129.62 = £76/\text{ha}$

$c_o = 2.2 \text{ x } 65.9 - 109.98 = £35/\text{ha}$

At the base year 300 ha of wheat and 200 ha of oat are grown. The auxiliary calibrated linear program for $\varepsilon = 0.01$ is

Max	Z	=		$76\ x_w$	+	$35\ x_o$	
s.t.							
Land	x_w		+	x_o	$\leq$	500	$[\lambda_{land}]$
Wheat calibration	x_w				$\leq$	300.01	$[\lambda_{wheat}]$
Oats calibration				x_o	$\leq$	200.01	$[\lambda_{oats}]$

At the optimal solution, the model, due to the calibration constraints, reproduces the base year situation and therefore $x^R = (x_w, x_o) = (300.01, 199.99)$, and the vector of dual variables is $\lambda^{cal} = (\lambda_{wheat}, \lambda_{oats}) = (41, 0)$. Applying the expressions (14) and (15) for $\boldsymbol{\alpha}$ and $\boldsymbol{\gamma}$ we take $\boldsymbol{\alpha} = (88.62, 110)$ and $\boldsymbol{\gamma} = (27.33, 0)$. The objective function now reads

$$f(x) = (2.98\text{x}69 - 88.62 + 0.5\text{x}27.33\, x_w)x_w, + (2.2\text{x}65.9 - 109.98)\, x_o$$

If the calibration constraints of the first stage are removed it turns out that the model with the nonlinear objective function calibrates exactly to the base year activity levels. Figure 3.1 illustrates graphically the calibration property of the nonlinear objective function.

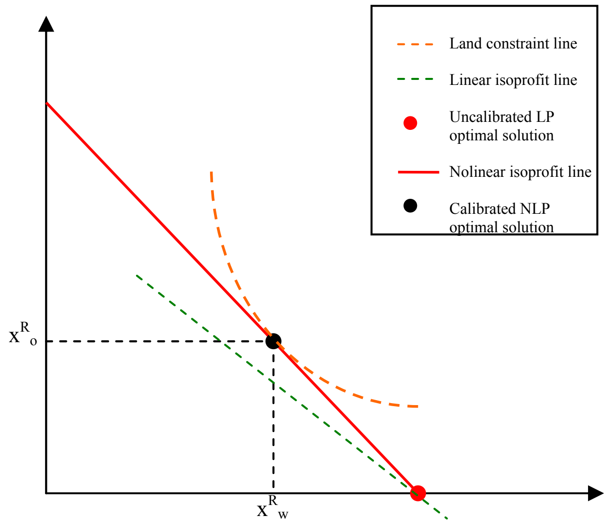

Figure 3.1. The PMP calibration of a MP model with two activities

3.2.2. Problems and Extensions of the Standard PMP Calibration Method

The fundamental assumption of PMP is that at least some production activities exhibit decreasing marginal returns. This can happen due to either decreasing yields or increasing variable marginal costs. The intuition of the PMP setup is that these variable marginal costs include accounting costs, *c* as well as the "hidden" or opportunity costs incorporated in the shadow values of calibrated activities, λ^{cal}. Paris and Howitt, (1998) use a specification of the total variable cost function C(x) being the integral of the vector of variable marginal costs ($\boldsymbol{\lambda}^{cal} + \mathbf{c}$) (derived at the first stage of PMP from LP problem 1) with respect to the vector $\mathbf{x}^R$ over the interval (0, $\mathbf{x}^R$). That reads

$$C(\mathbf{x}^R) = \int_0^{x^R} (\boldsymbol{\lambda}^{cal} + \mathbf{c})' dx = \boldsymbol{\lambda}^{cal\prime}\mathbf{x}^R + \mathbf{c}'\mathbf{x}^R$$

The choice of different functional specifications for the marginal cost corresponds to different but admissible specifications of the total variable function. For example, the result of postulating that

$$(\boldsymbol{\lambda}^{cal} + \mathbf{c}) = \mathbf{d} + \mathbf{Q}\mathbf{x}, \tag{16}$$

with $\mathbf{Q}$ a symmetric positive definite matrix of parameters associated with the quadratic term, is the familiar quadratic total variable cost function

$$C(\mathbf{x}) = \int_0^{x^R} (\lambda^{cal} + \mathbf{c})' dx = \mathbf{d}\mathbf{x}^R + \frac{1}{2}\mathbf{x}^{R\prime}\mathbf{Q}\mathbf{x}^R \tag{17}$$

Equation (16) indicates that the amount of information available for reconstructing the $\mathbf{Q}$ matrix is given by the vector ($\boldsymbol{\lambda}^{cal} + \mathbf{c}$) and the vector $\mathbf{x}^R$. When $\boldsymbol{J}$ is the number of observations in the data set for these two vectors, the number of parameters of the $\mathbf{d}$ vector, and of the $\mathbf{Q}$ matrix that must be recovered is $\boldsymbol{J}$ and [$\boldsymbol{J}(\boldsymbol{J}$ +1)/2] respectively. The problem of specifying $\boldsymbol{J}$ + [$\boldsymbol{J}(\boldsymbol{J}$ +1)/2] parameters on the basis of 2$\boldsymbol{J}$ pieces of information is usually solved by letting $\mathbf{d} = \mathbf{c}$ and setting all off-diagonal elements of $\mathbf{Q}$ to zero. The $\boldsymbol{J}$ diagonal elements of $\mathbf{Q}$, q_{ii}, can then be calculated as

$$q_{ii} = \frac{\lambda_i^{cal}}{x_i^R} \quad \text{for all i = 1,…,}\boldsymbol{J} \tag{18}$$

However, Heckelei, (1997), Henry de Frahan (2005) and Wiborg *et al*, (2005) point out that the above specification results in a potential problem because the model's behaviour is ultimately determined by this rather arbitrary specification of the employed objective function. To illustrate this, Heckelei used the supply function (marginal cost) for one of the preferable activities x^p after the PMP calibration:

$$mc_i = p_i = c_i + q_{ii}\, x_i^p + \mathbf{A}_i\, \lambda^{res} \tag{19}$$

Change in marginal cost by expanding or reducing x_i is determined by the parameter q_{ii}. This means for the response behaviour that, for example, when p_i or c_i are changed, the extent of change in the level of x_i is smaller the higher q_{ii} is and vice versa. Furthermore, this specification rule leads to a cost function which is linear in "marginal" activity levels, because the elements of $\boldsymbol{\lambda}^{cal}$ are zero. This in turn implies that $\boldsymbol{\lambda}^{res}$ remains constant, because it is determined by the profitability of the "marginal" activities alone which is constant per activity unit. Consequently, a price increase for products of the "preferable" production activities leads to a substitution of marginal activities, but leaves the other preferable activity levels unchanged until the first marginal activity is driven out of the basis.

As Heckelei (1997) suggests, to alleviate the above problem a higher q_{ii} can be obtained by setting the linear cost term d_i to zero in addition to the off diagonal elements of **Q** and calculating the diagonal elements q_{ii} as $(c_i + \lambda_i) / x_i^R$ without changing the calibration property. However, both analytical solutions require that the quadratic cost matrix **Q** is specified as strictly diagonal. A further weakness of the approach according to Bauer and Kansakoglu (1990) is that the costs implied in the non-linear part cannot explicitly be attributed to specific production factors.

Heckelei and Wolff (2003) and Heckelei (2005) have pointed out a third potential problem with the standard PMP approach. In the short review and critique on PMP within that paper the author shows that the shadow prices of resource constraints and the marginal variable costs at observed quantities enforced by the PMP approach are generally incompatible with the marginal conditions of the nonlinear model to be specified. He takes as an example the quadratic objective function

$$f(x) = \mathbf{p}'\mathbf{x} - \mathbf{d}'\mathbf{x} - \frac{1}{2}\mathbf{x}'\mathbf{Q}\mathbf{x}$$

or if letting $\mathbf{d} = \mathbf{c}$ and $\mathbf{g} = \mathbf{p} - \mathbf{c}$

$$f(x) = \mathbf{g}'\mathbf{x} - \frac{1}{2}\mathbf{x}'\mathbf{Q}\mathbf{x} \tag{20}$$

the optimisation problem becomes: max $f(x)$ subject to $\mathbf{Ax} \leq \mathbf{b}$ $[\boldsymbol{\lambda}^{res(Q)}]$, $\mathbf{x} \geq \mathbf{0}$ with the Lagrangian formulation

$$L(x) = = \mathbf{g}'\mathbf{x} - \frac{1}{2}\mathbf{x}'\mathbf{Q}\mathbf{x} + \boldsymbol{\lambda}^{res(Q)} [\mathbf{b} - \mathbf{Ax}] \tag{21}$$

If all optimal activity levels are positive the first order conditions in gradient format are obtained as

$$\frac{\partial L}{\partial x} = \boldsymbol{g} - \boldsymbol{Qx}^{\mathrm{R}} - \boldsymbol{A'\lambda}^{\mathrm{res(Q)}} = \boldsymbol{0} \tag{22}$$

$$\frac{\partial L}{\partial \lambda^{res(Q)}} = \boldsymbol{b} - \boldsymbol{Ax}^{\mathrm{R}} = \boldsymbol{0} \tag{23}$$

Solving (22) for **x** yields

$$\boldsymbol{x}^{\mathrm{R}} = \boldsymbol{Q}^{-1}(\boldsymbol{p} - \boldsymbol{d} - \boldsymbol{A'\lambda}^{\mathrm{res(Q)}}) \tag{24}$$

and λ^{res} (at optimum) is obtained by substituting the right hand side of (24) into (23)

$$\boldsymbol{\lambda}^{\mathrm{res(Q)}} = (\boldsymbol{A}\,\boldsymbol{Q}^{-1}\boldsymbol{A})^{-1}(\boldsymbol{AQ}^{-1}\boldsymbol{g} - \boldsymbol{b}) \tag{25}$$

Heckelei and Wolff (2003) conclude that the value of $\boldsymbol{\lambda}^{\mathrm{res(LP)}}$ calibrated by step 1 of PMP (equation (6b): $\boldsymbol{\lambda}^{\mathrm{res}} = \mathbf{g}_{\mathrm{m}}\ (\mathrm{A}^{\mathrm{m}})^{-1}$) is expected to be different from the one implied by the nonlinear model which is assumed to represent farmer behaviour (equation (25)). Hence, they suggest that the second stage of the standard PMP uses these “apparently” wrong values at the observed activity levels through enforcement of the marginal cost equations, thereby implicitly imposing biased values for the estimation of the marginal cost (equation (16)) as well. However, Howitt (2005a) shows that the shadow values derived from the linear model are identical to those from the resulting PMP model.

It suffices to show that the dual values $\boldsymbol{\lambda}^{\mathrm{res(Q)}}$ that are obtained from the quadratic model (or in fact from any other suitable nonlinear model) are the same as the dual values, $\boldsymbol{\lambda}^{\mathrm{res(LP)}}$ derived from the LP model of the first stage of PMP. Rearranging the relevant equations (6a): $\boldsymbol{\lambda}^{\mathrm{res(LP)}} = \mathbf{g}_{\mathrm{m}}\ (\mathrm{A}^{\mathrm{m}})^{-1}$and (6b): $\boldsymbol{\lambda}_{\mathrm{p}}^{\mathrm{cal}} = \mathbf{g}_{\mathrm{p}} - \mathbf{A}^{\mathrm{p}\prime}\boldsymbol{\lambda}^{\mathrm{res(LP)}}$ yields equations

$$\mathrm{A}^{\mathrm{m}\prime}\boldsymbol{\lambda}^{\mathrm{res(LP)}} = \mathbf{g}_{\mathrm{m}} \tag{26a}$$

$$\mathbf{A}^{\mathrm{p}\prime}\boldsymbol{\lambda}^{\mathrm{res(LP)}} = \mathbf{g}_{\mathrm{p}} - \boldsymbol{\lambda}_{\mathrm{p}}^{\mathrm{cal}} \tag{26b}$$

Since for the “marginal” activities $\boldsymbol{\lambda}_{\mathrm{m}}^{\mathrm{cal}} = 0$[4] the above expressions can be rewritten as

$$\mathrm{A}^{\mathrm{m}\prime}\boldsymbol{\lambda}^{\mathrm{res(LP)}} = \mathbf{g}_{\mathrm{m}} - \boldsymbol{\lambda}_{\mathrm{m}}^{\mathrm{cal}} \tag{27a}$$

$$\mathbf{A}^{\mathrm{p}\prime}\boldsymbol{\lambda}^{\mathrm{res(LP)}} = \mathbf{g}_{\mathrm{p}} - \boldsymbol{\lambda}_{\mathrm{p}}^{\mathrm{cal}} \tag{27b}$$

[4] This is a fundamental assumption of the original PMP method. For further information see Howitt (1995).

and in unpartitioned form as

$$\mathrm{A}'\boldsymbol{\lambda}^{\mathrm{res(LP)}} = \mathbf{g} - \boldsymbol{\lambda}^{\mathrm{cal}} \qquad (28)$$

From equation[5] (16) it is known that at optimum

$$\boldsymbol{\lambda}^{\mathrm{cal}} + \mathbf{c} = \mathbf{d} + \mathbf{Q}\mathbf{x}^{\mathrm{R}} \qquad (29)$$

Substituting the r.h.s from (29), and letting again **c = d** equation (28) becomes

$$\mathbf{A}'\boldsymbol{\lambda}^{\mathrm{res(LP)}} = \mathbf{g} - \mathbf{Q}\mathbf{x} \qquad (30)$$

which is the same as equation (22) obtained from the Heckelei and Wolff's quadratic model. Therefore, the dual values derived from the linear model of stage 1 for a given observation set are, in fact, numerically identical to the dual values that the quadratic model would generate for the same observations.

3.3. Discussion on PMP Calibration Approach

From equation (16) it can be observed that the shadow values of the nonlinear model, $\boldsymbol{\lambda}^{res}$, are determined by objective function entries and technological coefficients as well as the parameters of the **Q** matrix, q_{ii}. With respect to the calibration property of PMP, Heckelei and Wolff's (2003) argument about the incorrect shadow values seems irrelevant because the correct shadow values will be known only when the true q_{ii} parameters are used in the model. However, these true parameters are not known and should be estimated from sample data. With respect to the model's ability to calibrate to a baseline situation, when only one year is considered and relevant prior information is lacking there is no justifiable reason to choose any of the infinite sets of q_{ii} parameters that calibrate the model. The choice of the parameters q_{ii} that are derived from the first step of the standard PMP method can be seen simply as a way to get justifiable estimates. Nonetheless, the calibration property of those parameters does not necessarily mean that they are good estimates of the true parameters. They are just one of the infinite number of solutions that solve equation (16). It should be made clear that changes in the activity levels between sample observations will consequently generate different dual values to be used by the quadratic model at the

[5] This equation expresses the condition that the second stage nonlinear model uses the dual values $\boldsymbol{\lambda}^{\mathrm{cal}}$ whose value according to equation (6b) depends on $\boldsymbol{\lambda}^{\mathrm{res(LP)}}$, which Heckelei and Wolff (2003) argues are different from the $\boldsymbol{\lambda}^{\mathrm{res(Q)}}$values implied by the quadratic model.

second stage of PMP. Therefore an appropriate estimation procedure should be applied for deriving estimates of the true parameters possessing known statistical properties. The next section is concerned with this issue.

3.4. Estimation of PMP Model Parameters

First of all, as Howitt (2005b) suggests, there are two significant practical problems with the diagonal specification used in the standard PMP method in that it assumes that there are no "cross effects" between the amount of land allocated to, for example, crops, apart from the effect on the total land constraint. This assumption implies that there are no substitution or complementarity cost effects between production activities carried out in the same region or farm. However, the practice of rotations in crop production, for example, indicates that farmers are well aware of the interdependencies among crops and use them to stabilise or increase profits.

It is apparent that the assumption of diagonal **Q** is unrealistic, but to calibrate the full matrix of coefficients means that the number of unknown parameters of the full matrix, $[\boldsymbol{J}(\boldsymbol{J}+1)/2]$, is larger than the number of observations, $\boldsymbol{J}$ and the problem is said to be ill-posed. Therefore, there are infinite **Q** specifications that solve equation (16) and each one results in different simulation behaviour of the model. Fortunately, an approach based on information theory and the principle of maximum entropy (ME) is well suited to handle problems like this and to guarantee statistical properties of the chosen parameters.

3.4.1. The Maximum Entropy Estimation Method

Given non-experimental data generation processes in economics, many analytical models in economics contain estimates of unknowns that are unobserved and indeed are not accessible to direct measurement. In order to recover the unknown parameters, representing the economic system of interest, one is faced with an inverse problem that may be formalised in the following way:

$$\mathbf{y} = \boldsymbol{X}\beta \qquad (31)$$

where **y** is a T-dimensional vector of observations (data), β is an unobservable K-dimensional vector of unknowns and $\boldsymbol{X}$ is a known ($K \times T$) linear operator. The parameter of interest is defined as a solution of a functional equation depending on the data distribution. Hence the operator in the underlying inverse problem is in general unknown. The interest centres on an unknown and unobservable parameter, β which may be a number, a vector or a function. Without sufficient additional information (more observations) on $\boldsymbol{\beta}$ there exist an infinite number of $\boldsymbol{\beta}$ vectors that satisfy the model exactly.

3.4.1.1. The Maximum Entropy Principle

In information theory entropy is a measure for the amount of uncertainty or missing information in a signal or a data set. The more uncertainty ("chaos") exists the higher the entropy of a system. In information theory it is assumed that the information contained in an observation is inversely proportional to its probability. The –log(·) function (for an arbitrary log base) shown in Figure 3.2 is used to construct an information score based upon this description. Shannon (1948) employed an axiomatic approach to define a unique function to measure the uncertainty of a set of events. He defined the entropy of the distribution of probabilities, $\mathbf{p} = (p_1, p_2, \ldots, p_k)$, of a (discrete) random variable X, with possible observable outcome values X_k, $k = 1, 2, \ldots, K$, as the measure $H(\boldsymbol{p})$ where

$$H(\boldsymbol{p}) \equiv -\sum_{k=1}^{K} p_k \ln p_k \qquad (32)$$

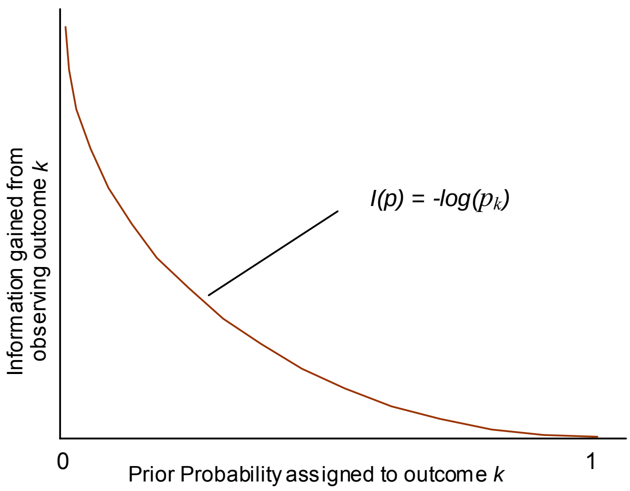

Figure 3.2. Information mapping under the –log(·) function

When $p_k = 1$ for one outcome and $p_k = 0$ for the rest, then because $p_k \ln(p_k)= 0$ $H(\mathbf{p})$ is minimised, whereas it is maximised when each event is equally likely *i.e.* $p_1 = p_2 = \ldots = p_k = 1 / K$ and has a unique solution. In other words, the expected information content from an event is greatest when the probabilities are uniform which follows from the 'Principle of Insufficient Reason'. The Shannon Entropy metric, $H(\mathbf{p})$, favours the set of relative frequencies that could have been generated in the greatest number of ways consistent with what is known.

Based on the work of Shannon, Jaynes (1957a, b) suggested that what is known (sample or non-sample information) can be viewed as constraints on the entropy measure, $H(\mathbf{p})$. Given that two or more feasible probability distributions that satisfy the constraints could be found, the one to be selected is the one which is least informative, or most uncertain according to the Shannon entropy criterion, namely the distribution with the highest (maximum) entropy. Put in another way, the distribution obtained by the maximum entropy (ME) crtierion 'agrees with what is known, but expresses maximum uncertainty with respect to all other matters', and therefore it is the "best" estimator. Therefore, Jaynes proposed the ME as a means by which to recover the unknown probabilities, **p** (or probability distribution) subject to prior information-related constraints.

The following example which is a variant of the dice problem, demonstrates intuitively how the ME method works. Suppose you are given a six-sided die that can take on the values $k = 1, 2, \ldots, 6$, and you are asked to estimate the probabilities $\mathbf{p} = (p_1, p_2, \ldots, p_6)'$ for each possible outcome in the next roll of the die. The only information that is given is the average outcome from a large number of independent rolls of the die, y. For a given mean value there are an infinite number of combinations of distributions supported on $\{1, 2, \ldots, 6\}$ (the number of combinations of the six probabilities) that could have generated the mean y. Note that the observed frequency distribution of the sample, which would yield the maximum likelihood estimator for a multinomial distribution, is not given. The problem is clearly ill-posed because there are six unknown probabilities but only one data point, the mean value, and the constraint that the probabilities have to add to one.

One way to solve ill-posed problems is by using prior or non-sample information to choose from the feasible set of solutions. The observed average draw of a 'fair' die for instance, is expected to match the mean of the discrete uniform distribution, y = 3.5. Then it can be asserted that the underlying probability distribution is discrete uniform because the sample information matches the prior belief. However, if $y \neq$ 3.5, the observations suggest that the die is not fair and the underlying distribution is not likely to be uniform. In situations like this, where the underlying distribution is not known, the ME formalism developed by Jaynes (1957a, b) comes to the rescue. Formally, the distribution of probabilities, $\mathbf{p} = (p_1, p_2, \ldots, p_6)'$ that maximises the entropy,

$$H(\mathbf{p}) = -\sum_{k=1}^{6} p_k \ln p_k \tag{33a}$$

subject to

$$\sum_{k=1}^{6} p_k x_k = y \tag{33b}$$

$$\sum_{k=1}^{6} p_k = 1 \tag{33c}$$

where $x_k = k$ for each k = 1, 2, …, 6 is most likely to be the probabilities underlying the die. Using a numerical optimisation package such as the Frontline Systems' **Solver**, the estimated entropy distributions, for various values of y, are plotted in Figure 3.4, and the entropy value for different y is plotted in Figure 3.3. As expected, y = 3.5 results in a discrete uniform distribution and a maximum value of $H(\mathbf{p})$.

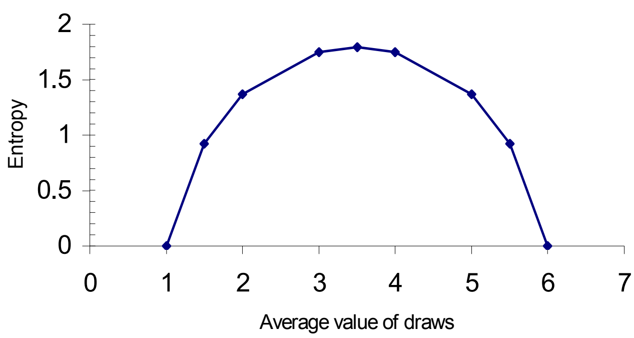

Figure 3.3. The entropy value for different average scores

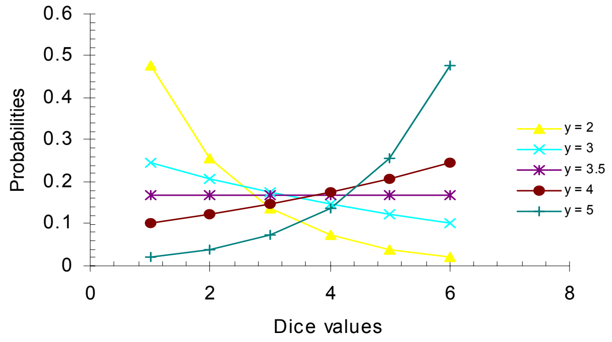

Figure 3.4. The estimated ME distributions for the die problem

3.4.2. The GME Estimation Method

Golan *et al.* (1996) generalised ME by expressing the unknown parameters and disturbances of the standard econometric problem (general linear model) in terms of discrete probability distributions.

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{e} \tag{34}$$

Where $\mathbf{y}$ is a T-vector of noisy observations, $\mathbf{X}$ is the ($T \times K$) design matrix composed of explanatory variables, $\boldsymbol{\beta}$ is a K-vector of unknown (unobservable) response parameters, and $\boldsymbol{e}$ is a T-vector of unobservable disturbances. Although $\boldsymbol{\beta}$ is taken to be a real-valued vector, Golan *et al.* treat each β_k as a discrete random variable belonging to a more specialised parameter space defined by a compact support set and $2 \leq M \leq \infty$ potential outcomes, $\mathbf{z_k} = [z_{1k}, z_{k2}, \ldots, z_{kM}]'$, with corresponding probabilities $\mathbf{p_k} = [p_{k1}, p_{k2}, \ldots, p_{kM}]'$. If z_{k1} and z_{kM} are the plausible extreme values (lower and upper bounds) of β_k, β_k can be expressed as a convex combination of these two points. That is, there exists $p_k \in [0, 1]$ such that, for $M = 2$,

$$\beta_k = p_k z_{k1} + (1 - p_k) z_{kM} \quad \forall \text{ k}$$

Similarly, Golan *et al.* represent the uncertainty about the outcome of the error process, reflected in the vector of disturbances $\boldsymbol{e}$, by treating each e_t as a finite and discrete random variable with $2 \leq J \leq \infty$ possible outcomes, $\mathbf{v}_t = [v_{t1}, v_{t2}, \ldots, v_{tJ}]'$, with corresponding probabilities $\mathbf{w}_t = [w_{t1}, w_{t2}, \ldots, w_{tJ}]'$. If there exist sets of error bounds, v_{t1} and v_{tJ}, for each e_t so that $\Pr[v_{t1} < e_t < v_{tJ}]$ may be made arbitrarily small. With $\mathbf{w}_t \in (0, 1)$, for $J = 2$, each disturbance may be written as

$$e_t = w_t v_{t1} + (1 - w_t) v_{tJ} \quad \forall\, t$$

If the error distribution is assumed symmetrical and centred about **0**, a symmetric support set, $v_{t1} = -v_{tJ}$, for each t can be specified.

Consistent with these specifications, the elements of $\boldsymbol{\beta}$ and $\boldsymbol{e}$ can be rewritten as

$$\beta_k = \mathbf{z}_k'\mathbf{p}_k = \sum_m z_{mk}\mathbf{p}_k \text{ for } k = 1, 2, \ldots, K,\ m = 1, 2, \ldots, M$$

$$e_t = \mathbf{v}_t'\mathbf{w}_t = \sum_j v_{tj}\mathbf{w}_t \text{ for } t = 1, 2, \ldots, T,\ j = 1, 2, \ldots, J$$

and in matrix form as

$$\boldsymbol{\beta} = \mathbf{Zp} = \begin{bmatrix} \mathbf{z}_1' & 0 & \cdots & 0 \\ 0 & \mathbf{z}_2' & \cdots & 0 \\ \vdots & \vdots & \cdots & 0 \\ 0 & 0 & \ldots & \mathbf{z}_k' \end{bmatrix} \begin{bmatrix} \mathbf{p}_1 \\ \mathbf{p}_2 \\ \vdots \\ \mathbf{p}_K \end{bmatrix}$$

$$\boldsymbol{e} = \mathbf{Vw} = \begin{bmatrix} \mathbf{v}_1' & 0 & \cdots & 0 \\ 0 & \mathbf{v}_2' & \cdots & 0 \\ \vdots & \vdots & \cdots & 0 \\ 0 & 0 & \ldots & \mathbf{v}_T' \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_T \end{bmatrix}$$

Using the reparameterised unknowns, $\boldsymbol{\beta} = \mathbf{Zp}$ and $\boldsymbol{e} = \mathbf{Vw}$, Judge and Golan (1992) rewrite the GLM as

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{e} = \mathbf{X\,Zp} + \mathbf{Vw} \qquad (35)$$

For k observations the "support points (values)" together with their (so far unknown) probabilities allow the calculation of the expected values of the parameters, β_k and e_k, as

$$E(\beta_k) = \sum_{m=1}^{M} z_{mk} p_{mk} \ \forall\, \mathrm{k}$$

$$E(e_t) = \sum_{j=1}^{J} v_{jt} w_{jt} \ \forall\, t$$

which can be interpreted as the "estimates" of the parameters. The set of probabilities of the support points that add the least amount of information to be consistent with the data is found by using the ME metric in the following way:

$$\max H(\boldsymbol{p}) = -\sum_{m=1}^{M}\sum_{k=1}^{K} p_{km} \ln p_{km} - \sum_{j=1}^{J}\sum_{t=1}^{T} w_{tj} \ln w_{tj} \tag{36a}$$

$$\text{s.t. } \sum_{k=1}^{K} x_{kt} E(\beta_k) + E(e_t) = y_t \quad \forall\, t \tag{36b}$$

$$E(\beta_k) = \sum_{m=1}^{M} z_{mk} p_{mk} \ \forall \text{ k} \tag{36c}$$

$$E(e_t) = \sum_{j=1}^{J} v_{jt} w_{jt} \ \forall\, t \tag{36d}$$

$$\sum_{m=1}^{M} p_{mk} = 1 \quad \forall\, k \tag{36e}$$

$$\sum_{j=1}^{J} p_{jt} = 1 \quad \forall\, t \tag{36f}$$

The first set of constraints guarantees that the resulting mathematical expectations of the parameters satisfy exactly the t observations in the form of the GLM (data consistency). The second set of constraints defines the expected values of the parameters and the third express the summation condition of probabilities.

3.4.2.1. A Simple Example of Maximum Entropy Parameter Estimation[6]

Consider the simple quadratic cost function:

$$C = \alpha x + \frac{1}{2}\beta x^2$$

where x denotes output.

The problem is to estimate the two parameters, α and β when there is only one observation available in which the marginal cost is 40 when the output x = 10. The data relationship that has to be satisfied is:

$$40 = a + 10b$$

It is clear that there are an infinite number of parameter values (feasible set of probabilities) for α and β that satisfy this relationship. Suppose five discrete values for a support space. If negative costs are ruled out, the lower support space is bounded at zero. The upper extreme value can be defined by the coefficient value

[6] This is a variant of the example found in Howitt (2005b pp 163-4)

that would explain all of the cost when the other coefficient is zero. Using this as a basis for the support points, five evenly distributed support values would be:

$$z\alpha_i = [0, 8, 16, 32, 40] \quad \& \quad z\beta_i = [0, 1, 2, 3, 4]$$

There are a number of feasible sets of probabilities that would solve the following system of equations:

$$40 = \sum_i z\alpha_i p\alpha_i + 10\sum_i z\beta_i p\beta_i$$

$$\sum_i p\alpha_i = 1$$

$$\sum_i p\beta_i = 1$$

$$p\alpha_i, p\beta_i \geq 0$$

For all the different possible sets the resulting cost function C will be different and there is no reason to favour one from the others. The following ME problem solves for the two distributions that are most likely to have generated the total cost function when the marginal cost is 40 for an output $x = 10$:

$$\text{Max } H(\mathbf{p}) = -\sum_i p\alpha_i \ln p\alpha_i - \sum_i p\beta_i \ln p\beta_i$$

$$\text{Subject to} \quad 40 = \sum_i z\alpha_i p\alpha_i + 10\sum_i z\beta_i p\beta_i$$

$$\sum_i p\alpha_i = 1$$

$$\sum_i p\beta_i = 1$$

$$p\alpha_i, p\beta_i \geq 0$$

The solution to this problem, namely the expected values of parameters α an β, is plotted on the histograms in the figures below.

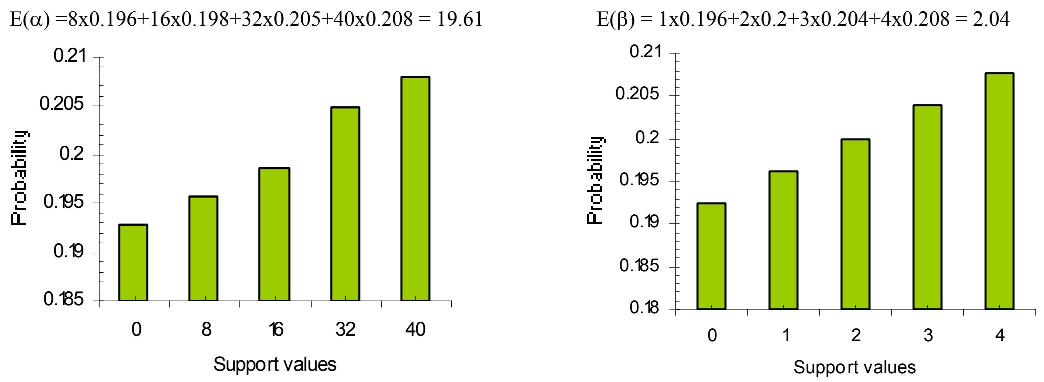

Figure 3.5. The maximum entropy solution for the parameters in the form of probability distributions

3.5. Maximum Entropy and PMP

The number of applications and the use of ME in agricultural economics are limited but they provide an insight into its potential. The study of Miller and Lence (1998) deals with the lack of activity-specific input data when analysing multiproduct-multifactor firms. They use a GME approach to estimate multi-output production functions with incomplete data. The paper by Léon et al. (1999) proposes the use of GME method to estimate input-outpout coefficients, which reflect the unobserved allocation of farm input accounting costs to the various outputs produced. Fraser (2000) considers the case of demand functions for meat in UK to assess the ease of implementation of GME and identify strengths and weaknesses of the approach as an applied estimation technique. In an effort to show that there is a continuum of analysis between mathematical programming and traditional econometric techniques Paris and Howitt' (1998) demonstrate how to recover flexible cost functions from very limited data sets using a ME approach. Heckelei and Britz (2000) further expand that approach in order to exploit information contained in a cross sectional sample for the specification of nonlinear cost functions in regional programming models. In a more recent study, Howitt and Reynaud (2003) use ME for a spatial dissagregation of agricultural production data. The rest of this section presents in more detail how ME can be applied in a PMP context.

On a regional level, a typical data set of multi-input/multi-output production is usually restricted to cropland allocations by production activities, the total output of the various activities, and their prices. Information on the crop marginal production costs is usually absent or less accurate. This is particularly true with micro data on land class variability, technology and risk. Apparently, this information often features in the farmers' decisions, but is unavailable to the model builder. The technological, market and environmental constraints determining the opportunity costs the farmer faces may not be revealed explicitly by the sample information but are assumed to be reflected in the marginal crop and livestock allocation decisions he/she takes. Hence, the vector of farm outputs $\mathbf{x}^R$ must incorporate and reflect information about costs as perceived by the farm entrepreneur. It is the task of the modeller to extract the maximum amount of economic information from these incomplete data, to decode the encrypted cost information contained in the $\mathbf{x}^R$ vector, and to reconstruct a total variable cost function in a way suitable for policy analysis.

If the following quadratic function

$$C(x) = \frac{1}{2}\,\mathbf{x}'\mathbf{Q}\mathbf{x} \tag{37}$$

is assumed to describe the total variable cost, with $\mathbf{Q}$ a symmetric positive definite matrix of parameters associated with the quadratic term, the marginal cost corresponds to $\mathbf{Qx}$. As section 3.3.2 shows, PMP rationale assumes that

$$\mathbf{Qx}^R = (\boldsymbol{\lambda}^{cal(LP)} + \mathbf{c}) = \mathbf{mc} \tag{38}$$

which encapsulates the available amount of information for estimating the $\mathbf{Q}$ matrix. As already noted there, when $\boldsymbol{J}$ is the number of observations in the data set (*i.e.* R = $\boldsymbol{J}$) for these two vectors, the number of parameters of the $\hat{\mathbf{Q}}$ (estimated **Q)** matrix that must be recovered is [$\boldsymbol{J}(\boldsymbol{J}$ +1)/2]. The problem of specifying [$\boldsymbol{J}(\boldsymbol{J}$ +1)/2] parameters on the basis of $\boldsymbol{J}$ pieces of information is solvable by the ME method. The second order conditions require that the Hessian[7] matrix of the cost function is positive definite for the resulting matrix PMP model (max$\{\mathbf{g}'\mathbf{x} - 1/2x'\hat{\mathbf{Q}}\mathbf{x}\}$ s.t. $\mathbf{Ax} \leq 0$) to converge to a stable solution. In order to ensure that the matrix $\mathbf{Q}$ is symmetric definite positive, Paris and Howitt (1998) suggest the implementation of the Cholesky factorisation:

[7] The Hessian of a function $f(x_1, x_2, \ldots, x_n)$ is the n x n matrix whose ijth entry is $\partial^2 f / \partial x_i \partial x_j$

$$\mathbf{Q} = \mathbf{LDL'}, \tag{39}$$

where **L** is a unit lower triangular matrix, **L′** is its transpose and **D** is a diagonal matrix. In order to cast the problem of recovering the **Q** matrix in a ME framework using the Cholesky factorisation, Paris and Howitt (1998) regard each parameter of the **L** and **D** matrices as the expected value of an associated probability distribution defined over a set of known discrete support values. Each non zero element of the **L** matrix is defined as $E(l_{jj'})$, and $p(l_{jj',k})$ is the probability that the element $l_{jj'}$ takes the (support) value $l_{jj',k}$. Thus, $E(l_{jj'})$ is given by the expression:

$$E(l_{jj'}) = \sum_{k=1}^{r} l_{\mathrm{jj'},\mathrm{k}}\, p(l_{jj',k}) \text{ for all } j,j' = 1, 2, \ldots, J \;\; k = 1,\ldots,K$$

or for the matrices $\boldsymbol{L} = \boldsymbol{Z}^{\mathrm{L}}\boldsymbol{P}^{\mathrm{L}}$

where r is the number of all possible discrete values of the element $l_{jj'}$, $\boldsymbol{Z}^{\mathrm{L}}$ is the known support-space matrices for the probability distributions of the $\boldsymbol{L}$ matrix and $\boldsymbol{P}^{\mathrm{L}}$ is the probability matrix. The elements of the matrix **D**, $E(d_{jj'})$, are positive and determined similarly by the following expression:

$$E(d_{jj'}) = \sum_{k=1}^{r} d_{\mathrm{jj},\mathrm{k}}\, P(d_{jj,k}) \;\; k = 1,\ldots,K$$

and $\boldsymbol{D} = \boldsymbol{Z}^{\mathrm{D}}\boldsymbol{P}^{\mathrm{D}}$

In general, the specification of suitable support range for $\boldsymbol{L}$ and $\boldsymbol{D}$ is easily defined using the marginal cost values. The selection of the support intervals constitutes the only subjective aspect of the approach however, (Paris & Howitt, 1998) give a policy analysis example which shows that the response of the model is robust with respect to widely different specifications of these intervals. Usually, the support values of $l_{jj'}$, $l_{jj',k}$, are set out from the base value of λ_j^{cal} / x_j, which when multiplied for r equidistant values, yields the grid of possible values. Note that the base values correspond to q_{ij} elements when **Q** is a diagonal matrix. By proceeding in a similar way, the possible values of d_{jj} are determined. The maximum entropy problem (problem 4.1) consists in recovering the $\boldsymbol{p}^{\mathrm{L}}$ and $\boldsymbol{p}^{\mathrm{D}}$ such that

$$\max H(\boldsymbol{p}^{\mathrm{L}}, \boldsymbol{p}^{\mathrm{D}}) = - \sum_{jj',k} \mathbf{p}^{L} \log \mathbf{p}^{L} - \sum_{jj,k} \mathbf{p}^{D} \log \mathbf{p}^{D} \tag{40a}$$

subject to

$$\mathbf{mc} = \mathbf{Qx}^{R} = \mathbf{LDL'x}^{R} = (\boldsymbol{Z}^{L}\boldsymbol{p}^{L})(\boldsymbol{Z}^{D}\boldsymbol{Z}^{D})(\boldsymbol{Z}^{L}\boldsymbol{p}^{L})'\mathbf{x}^{R} \tag{40b}$$

$$1 = \sum_{k} p_{k}^{L} \tag{40c}$$

$$1 = \sum_{k} p_{k}^{D} \tag{40d}$$

where the last two constraints express the summation condition of probabilities.

The entropy criterion in the objective function equation looks for the set of probabilities which adds the least amount of information *i.e.* deviates the least from a uniform distribution over the support points, but satisfies the explicitly shown "data constraints" (40b)-(40d) of the estimation problem being the marginal cost condition, and the summation condition of probabilities. This objective (of requiring the weakest assumptions about the model) is achieved by using the postulate that every desired parameter is the product of a probability distribution and its corresponding set of known support values (Paris and Howitt, 1998).

By adopting this feasible and consistent approach, it is possible for the modeller to exploit all the available information to the maximum extent (Paris & Howitt, 1998). Since empirical data always contain some amount of information, even when an estimation problem is ill-posed, in the absence of any better source of such information, use of the ME approach allows modellers to satisfy the dictum that information should never be discarded.

3.6. Discussion and conclusion

The most important property of a programming model for agricultural supply is its simulation behaviour; *i.e.* whether it is capable of capturing the behavioural response of producers to changing economic conditions, so that it is capable of evaluating impacts of political, market or technical developments relevant to agriculture. The PMP approach provides an elegant way to calibrate programming models to observed behaviour and renders a more realistic and smooth aggregate supply response relative to a linear programming model. Models that are capable of reproducing base year observations are more likely to capture behavioural responses to changing economic incentives than models that are not capable of reproducing base year observations. These merits have led to a widespread application of PMP approaches in the context of aggregate agricultural programming models.

However, to date only Heckelei and Wolff (2003) have addressed the question of whether the PMP procedure itself is designed to make best use of additional data information. They look at the information from an econometrician's point of view and points out that there is a fundamental problem of the PMP methodology in the context of simulations with agricultural sector models. The shadow prices, λ^{cal}, actually do not only capture the hidden costs of production, but, rather, also any type of model mis-specification, such as data errors, aggregation bias, wrong or lacking representation of risk behaviour etc. Also, the contribution of some of these misspecifications might depend on particular economic and policy scenarios.

It should be made clear that PMP's calibration property alone does not necessarily imply any quantitative realism to the response behaviour of the PMP-calibrated model. For example, in the case of the quadratic specifications there are an infinite number of parameter sets which satisfy the specification conditions, *i.e.* lead to a perfectly calibrated model, but each set implies a different response behaviour to the ever changing economic conditions. One problem is the thin information base provided by just one year of observations on activity levels, which does not provide any information on second order properties (the Hessian matrix) of the nonlinear objective function (Heckelei and Wolff, 2003). The different methods developed to choose among the infinite number of calibrating parameter sets increasingly recognise the need to introduce additional information in order to avoid arbitrary simulation behaviour.

The employment of the Maximum Entropy criterion in PMP methodological procedure, which generally allows the use of more than one observation on activity levels, introduced an econometric criterion for the specification of PMP models and was the first step to bridge the traditional gap between econometric and optimisation models (Paris and Howitt, 1998; Heckelei and Wolff, 2003). By incorporating more than one observation the ME approach can provide a generally applicable tool for estimation of the model parameters underlying the observed response behaviour of farmers.

4. The Scottish Agricultural Systems Economic Model (SASEM)

4.1. Introduction

The previous chapter showed how total variable cost functions can be estimated using the ME and GME approaches. The principal message of that chapter is that empirical data always contain a certain amount of information and that, in the absence of a better source, such information should be used. The total amount of information contained in a poor sample of data generated by some phenomenon which gives an ‘out of focus’ picture of it, can be rearranged (by maximum entropy) to extract a better picture. One source of data for Scottish farms is the June Agricultural Survey. It is an annual census of agricultural activity which collects information relating to land use, crops, livestock, labour, horticulture and glasshouse production. Another source of secondary data is the Farm Management Handbook published by the Scottish Agricultural College, which is a reference to accounting and management data representative of “typical” farm businesses in the UK. The aim of this chapter is twofold. One objective is to present the SASEM model of Scottish agriculture that has been built as part of the bi-level programming model developed for agri-environmental policy optimisation. The other objective is to examine whether combined data from the above mentioned sources are sufficient to construct a simple aggregate, national-scale, agricultural production and land use change model by using the PMP rationale and ME-based estimation approaches.

The next section explains how the PMP rationale is suitable for constructing aggregate models. Following this, a short description and specification of the national scale macro-farm model of Scottish farming systems is given. The calibration and estimation procedures are then presented together with the corresponding results. The following section validates the simulation behaviour of the resulting models and the chapter concludes with a discussion on the application of the methods and the main points deriving from the findings.

4.2. PMP and Aggregation

PMP development was originally intended to calibrate agricultural economic production models. It was almost immediately used by a growing number of models at a regional (*e.g.*, Arfini and Paris, 1995; Judez *et al,* 2001; Judez *et al*, 2002; Umstätter, 1999) and sectoral level (*e.g.* Barkaoui and Butault, 2000; Heckelei and Britz, 2000). The present study, for reasons explained in section 4.3, required the construction of a national-scale model. What follows aims to explain why PMP-calibrated models offer a better alternative than LP models in this context.

As discussed in Chapter 2, in the transition from farm-level to regional-level analysis, ideally, a LP model should be constructed for every individual farm and all the individual models linked together to form the aggregate model. Since that approach is not feasible, sector models must either be based on representative farms, or on aggregate regional type models. It was pointed out in Chapter 2 that the latter approach involves aggregating the resources of a homogeneous region or area and modelling these aggregated data as a single large farm. An important drawback to this approach, *aggregation bias*, was identified in that discussion. The nature of aggregation bias is illustrated in the example given by Hazel and Norton (1986 pp 144) which is reproduced here. Consider the following two farm LP problems, each with two cropping activities, the first x_1 and x_2, and the second x_2 and x_3.

Farm A			RHS
Max Profit:	$60x_1$	$90x_2$	
Resource 1:	$1x_1$	$2x_2$	≤ 5
Resource 2:	$1x_1$	$1x_2$	≤ 5

Farm B			RHS
Max Profit:	$90x_2$	$100x_3$	
Resource 1:	$2x_2$	$1x_3$	≤ 10
Resource 2:	$1x_2$	$3x_3$	≤ 10

The optimal solution for farm A is to grow 5 units of x_1, while for farm B is to grow 5 units of x_2. The problem of the aggregate farm that represents both farms in a sector model would be as follows:

Aggregate Farm				RHS
Max Profit:	$60x_1$	$90x_2$	$100x_3$	
Resource 1:	$1x_1$	$2x_2$	$1x_3$	≤ 15
Resource 2:	$1x_1$	$1x_2$	$3x_3$	≤ 15

The optimal solution for the aggregate farm is to grow 15 units of x_1, which happens to be different from the sum of the solutions obtained from the individual farm models. By enabling farms to combine resources in proportions not available to them individually, the aggregate farm model imposes aggregation bias. As already discussed in Chapters 2 and 3 one form of such bias is the *overspecialisation* which is demonstrated in the above example by the result that the two farms separately grow two different crops, whereas when aggregated only one crop is included in the optimal solution. Below it is shown how aggregation bias can be minimised with the PMP approach.

The agricultural sector consists of a large number of farm businesses (some of which produce similar outputs), each one operating on land of different quality, facing different gross margins per unit of activity (*i.e.* marginal gross margins), g_i, using different technology and with different resource endowments. Suppose a region consists of nine arable farms all of which produce at least one common crop *c* among them. When, for that activity, marginal gross margins are constant (indicated by continuous horizontal lines in graph 4.1), a calibration constraint (indicated by the dashed vertical lines in Figure 4.1) is needed to limit the level of the activity to its observed value x_i as shown on the *x*-axis. The length of first segment, x_1, shows the level of the activity observed in the first farm, the length of the second segment, x_2, shows the level of the activity observed in the second farm and so on. Thus, the sum of all segments gives the level of that activity observed in the whole region.

If one follows the aggregate farm approach and treats the region as a single-macro farm, some sort of average values for gross margins (average technology is also factored therein) and resource availability, which represent the whole region, for every activity has to be used. This approach, as shown above, causes aggregation bias unless a set of very strong assumptions and conditions suggested by Day (1963) (see Chapter 2: 2.4) are met.

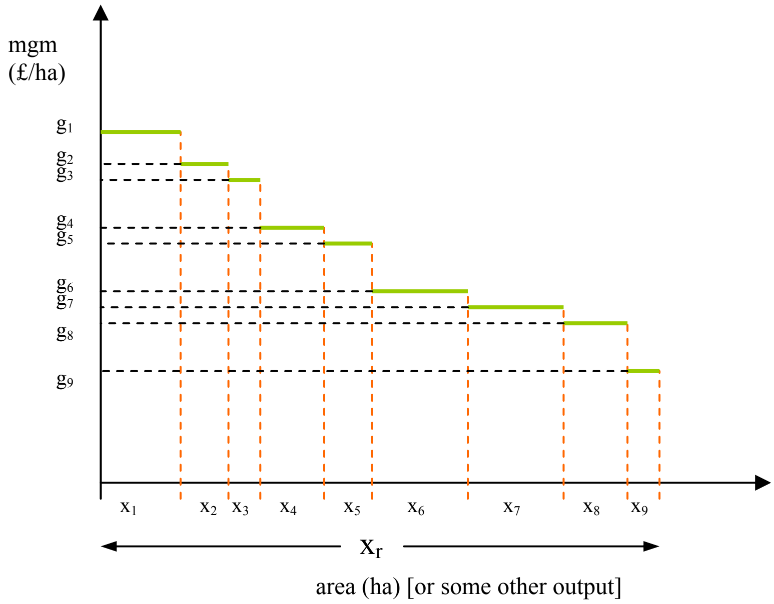

Figure 4.1. Relationship between marginal gross margins of an activity and its aggregate supply level

When these conditions cannot be satisfied, PMP, due to the non-linearity of the objective function that it generates, can offer an alternative with no additional requirements. As can be seen from Figure 4.1 the marginal gross margins of an activity for the whole region is likely to follow a step-wise downward line in relation to its level. This is because different farms undertake the same activity facing different (but assumed constant) gross margins per unit of activity. In fact, the shape of this step-wise curve reveals the structure of the industry (in the sense of a specific production activity) in a particular region or in a whole country *i.e.* distribution of farms of different type (reflected in their gross margins) with respect to the supply of a specific activity. Figure 4.2 shows four different possible structures for the nine arable farms region: (a) linear where the number of farms with high, intermediate and low marginal gross margins is similar; (b) convex where the number of farms in any marginal gross margins range is analogous to the marginal gross margins; (c) concave where the number of farms in any marginal gross margins range is reversely analogous to the marginal gross margins; (d) sigmoid where there are more farms with intermediate gross margins and less farms with high and low gross margins.

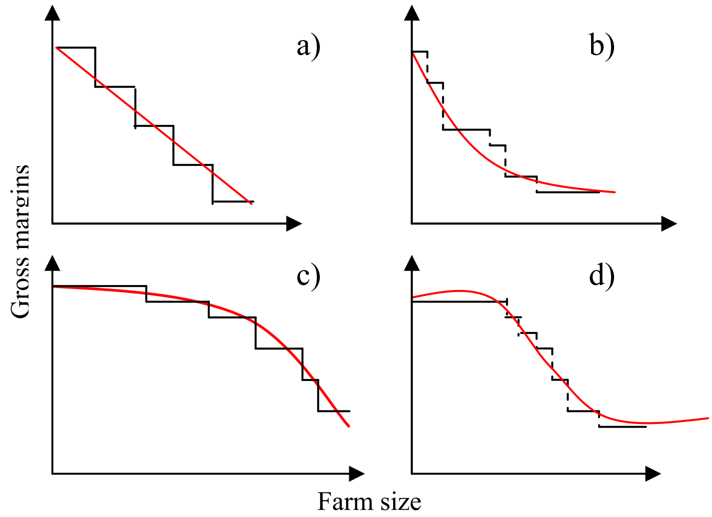

Figure 4.2. Characterisation of a sector based on the distribution of farms with respect to their marginal gross margins a) linear, b) convex, c) concave, d) sigmoid.

For each activity any of the above cases are possible to be observed which means effectively different marginal gross margin functions. Nonetheless, when data are not available, all cases can be approximated better by linear (case (a) in Figure 4.2) than by constant gross margins. For example, if a quadratic cost function is assumed the costs are increasing linearly (dashed upward-sloping line in Figure 4.3) with output and the opposite is true for the marginal gross margins. At optimum, the PMP-calibrated model solution for each and every calibrated activity will be the same as the activity's aggregate level observed in the base year.

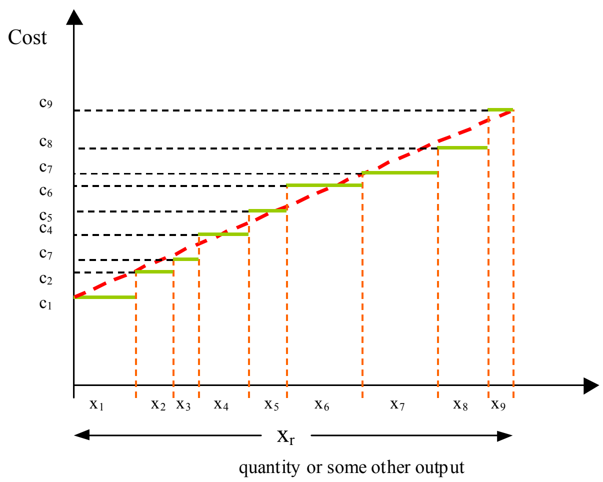

Figure 4.3. Relationship between variable costs of an activity and its aggregate supply level in case of a quadratic total variable costs function

Farms that are assumed to that have constant marginal gross margins for their activities can be modelled by LP. However, as this section showed, even when constant marginal gross margins are assumed for all farms in a region, the aggregate observed level of any activity is better approximated when a PMP-calibrated model with nonlinear total variable cost function is used. This advantage makes PMP-calibrated aggregate model more suitable than LP ones.

4.3. Specification of the Model

The basic premise underlying PMP is that allocation of agricultural land to various production activities is driven primarily by the responses of individual farmers to economic forces (for example, the relative profitability of crops and costs of factors and inputs). Given that the farmers' model will be used in the Follower Module of the bi-level programming model, the objective is to produce quantitative forecasts of agricultural land use and associated socio-economic and environmental impact, under different scenarios. In order to be able to predict the outcomes of hypothetical policy scenarios, the economic model must be sufficiently flexible to cope with a wide range of policy instruments. The model should be capable of including a wide range of different production activities. Production responses must be measured both in economic and physical terms since it is changes in, for example, crop areas or livestock numbers rather than revenues and costs that are required as input to the objective functions of the policy model (or to ecological & hydrological models that would give the values of the indicators which would then be used by the policy model). The structure of the economic model must also permit easy transfer of input and output data to and from the policy model. Finally, and most importantly, the model should demonstrate an ability to produce output close to observed situations.

The model is a comparative statics and partial equilibrium model in the sense that it ignores effects in and by other industries in the economy, and therefore, market prices of production factors, intermediate inputs and commodities are determined exogenously. It focuses on national production of the main cropping and livestock activities and it treats Scotland's farming industry as if it were a single farm. This choice was based on three reasons: a) agricultural policy objectives and measures are usually national in scope, b) census data needed for the PMP model is available nationally and, c) simplicity of analysis.

The following sections describe the construction of the economic model. Specifically, information on its spatio-temporal resolution and the range of activities included, as well as on its mathematical specification is provided. Following this description its calibration, estimation and validation are presented.

4.3.1. Spatial and Temporal Resolution

In a policy analysis context the relevant spatial scales range from individual plants and animals to global ecologies. Temporal scales range from hours to decades. Here, it is assumed that policy planning is concerned with determining the 'best' mix of policy instruments for the promotion of sustainable and multifuncional agroecosystems at a national scale. The choice of spatio-temporal scale for modelling depends on two main factors: The purpose of the analysis and the availability of data. Since the aim of the present modelling exercise is to support decisions at the national rather than the local level, the policy planning domain is chosen as the scale of the model. Figure 4.4 shows how this domain fits with other land-use planning domains characterised by their spatial and temporal scales. Although the temporal scale ranges from one to several years (a legislative term for example), the model only compares two different (partial) equilibrium states, before and after a change in one of the policy variables. Being a static model it does not study the motion towards equilibrium, or the process of the change itself.

From the choice of scale, it follows that farming systems should be modelled in some aggregate manner. It has been shown in the preceding section that one form of aggregation is to model all farms together as if they were a single-macro farm. Although this overstates flexibility and co-ordination of agricultural market, it is a widely accepted means of modelling large areas (see Moxey et al, 1995; Jones *et al*, 1995). Availability of time series of output data for the whole of Scotland which are suitable for applying the GME method was also an important motivating factor for selecting a national scale model.

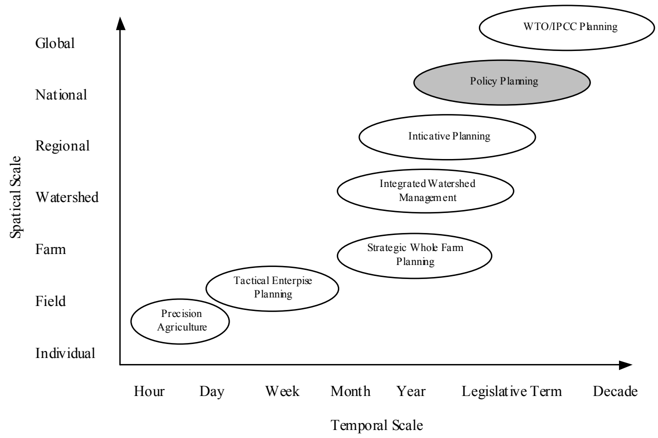

Figure 4.4. Spatio-temporal land-use planning scales (adopted from Mathews, 2001 pp 2)

4.3.2. Production activities covered in the Model

The choice of activities to be included in the model was based upon historical livestock numbers and land uses occupying large areas of land across Scotland, as reported in the June agricultural census. The aim of the model is to support decisions at the national rather than the local level and consequently small-scale activities, for example poultry and pig rearing, and minor more specialised activities that are restricted to a few farms were disregarded. Those selected are presented in Table 4.1. Decisions of the levels of different activities are expressed on a per hectare basis in the case of crops and in terms of animal numbers in case of livestock activities.

The dairy, beef and sheep sectors are represented in the model by the main livestock breeding and production activities practiced in Scotland. In this way, inter-farm links, such as lambs born on upland farms being sold to lowland farms for fattening, and dairy calves being sold to beef finishers, as well as potential knock-on effects between businesses within the two sectors can be analysed. Tables 4.2 and 4.3 show the balances between cattle activities and between sheep activities respectively as they are incorporated in the PMP model.

Table 4.1. Activities included in the model

Cash Crops	Forage Crops and Inputs	Cattle	Sheep
1. Winter wheat	1. Permanent Grassland for grazing	1. Dairy Cows (Spring/Sum Calving) 1a. Overwinter Cows 1b. Intensive Finish Calves	1. Hill Breeding Ewes for Store Lamb Production 1a. Sell Ewes to Upland for Breeding 1b. Store Lambs
2. Winter barley	2. Rotational Grassland for grazing	2. Dairy Cows (Autumn Calving)	2. Hill Ewes for Finished Lamb & Gimmer Production 2a. Finish Lambs 2b. Sell for Gimmering 2c. Store Lambs
3. Spring barley	3. Permanent grassland for Hay	3. Hill Suckler Cows 3a. Overwinter Calves 3b. Intensive Finish Calves	3. Upland Breeding Ewes 3a. Sell for Gimmering 3b. Store lambs 3c. Finish Lambs
4. Winter oats	4. Rotational grassland for Hay	4. Upland / Lowground Cattle (Spring/Summer calving, Silage diets) 4a. Overwinter Calves 4b. Intensive Finish Calves	4. Lowground Breeding Ewes for Early Finished Lamb 4a. Draft Ewes 4b. Finish Lambs
5. Spring oats	5. Permanent grassland for Silage	5. Upland / Lowground Cattle (Autumn calving, Silage diets)	5. Draft Ewes for Finish and Store Lambs 5a. Store Lambs 5b. Finish Lambs
6. Winter oilseed rape	6. Rotational grassland for Silage	6. Upland / Lowground Cattle (Spring/Summer calving, Straw diets) 6a. Overwinter Calves 6b. Intensive Finish Calves	6. Gimmerring
7. Spring oilseed rape	7. Turnips & Swedes	7. Upland / Lowground Cattle (Autumn calving, Straw diets)	7. Finish Lambs in Winter
8. Triticale	8. Rape	8. Overwintering Spring Calves	8. Finish Lambs in Autumn
9. Set-aside	_________________	9. Intensive Finishing Spring Calves at 12 months	
	1. N Feritliser	10. Finishing Autumn Calves	
	2. Sprays	11. Finishing Spring Calves at 18-24 months	
	3. Labour		

Spring/summer born dairy calves that are not intended for replacements can either be sold to finishers for intensive finishing or to other units for overwintering (balance rows 1, 5, 7) whereas autumn born dairy calves are only sold for finishing (balance row 6). It is assumed that 2/3 of the non replacement calves produced from hill suckler cows are born in spring-summer whereas the other 1/3 are born in the autumn. The former produce calves for sale to either overwintering units or directly to finishers (balance rows 2, 5, 7). The latter are sold directly for finishing. Upland and lowground cattle are divided in two groups depending on their diet, mainly silage or mainly straw. Each group is further divided in two groups depending on the time of calving, spring-summer or autumn. All spring-summer born non replacement calves are sold for fattening or finishing (balance rows 3, 4, 5, 7). Autumn born calves for sale are sold for finishing (balance row 6). Finally, all overwintered calves after fattening are finished (balance row 8).

Table 4.2. Cattle balance activities and constraints

Balance Rows	1. Dairy Cows (Spring-Summer Calving))	1a. Overwinter Dairy Calves (Spring-Summer Calving)	1b. Intensive Finish Dairy Calves	2. Dairy Cows (Autumn Calving)	3. Hill Suckler Cows	3a. Overwinter Hill Suckler Calves	3b. Intensive Finish Hill Suckler Calves	4. Upland/Lowground Cattle (Fb-Jn, Silage)	4a. Overwinter Upland/lowground Calves (Fb-Jn, Silage)	4b. Intensive finish Upland/Lowground Calves (Fb-Jn, Silage)	5. Upland/Lowground Cattle (Ag-Oc, Silage)	6. Upland/Lowground Cattle (Fb-Jn, Straw)	6a. Overwinter Upland/lowground Calves (Fb-Jn, Straw)	6b. Intensive Finish Upland/Lowground Calves (Fb-Jn, Straw)	7. Upland/Lowground Cattle (Ag-Oc, Straw)	8. Overwintering Spring Calves	9. Intensive Finishing Spring Calves at 12 months	10. Finishing Autumn Calves	11. Finishing Spring Calves at 18-24 months		RHS
1. Dairy Calves	-0.70	1	1																	≤	0
2. Hill Suckler Calves					-0.44	1	1													≤	0
3. Upland/lowground Calves Silage								-0.70	1	1										≤	0
4. Upland/lowground Calves Straw												-0.70	1	1						≤	0
5. Spring Calf Sales			-1				-1			-1				-1			1			≤	0
6. Autumn Calf Sales				-0.70	-0.26						-0.70				-0.70			1		≤	0
7. Overwitnering Spring Calves		-1				-1			-1				-1			1				≤	0
8. Finish Overw/ing Calves																-1			1	≤	0

Table 4.3. Sheep balance activities and constraints

Balance Rows	1. Hill Breeding Ewes for Store Lamb Production	1a. Sell Ewes to Upland for Breeding	1b. Store Lambs	2. Hill Ewes for Finished Lamb & Gimmer Production	2a. Finish Lambs	2b. Sell for Gimmering	2c. Store Lambs	3. Upland Breeding Ewes	3a. Sell for Gimmering	3b. Store lambs	3c. Finish Lambs	4. Lowground Breeding Ewes for Early Finished Lamb	4a. Draft Ewes	4b. Finish Lambs	5. Draft Ewes for Finish and Store Lambs	5a. Store Lambs	5b. Finish Lambs	6. Gimmerring	7. Finish Lambs in Winter	8. Finish Lambs in Autumn		RH S
1. Hill store lamb production	-0.75	1	1																		≤	0
2. Hill to Upland Breeding		-1						0.23													≤	0
3. Upland Breeding								-1.3	1	1	1										≤	0
4. Hill finished lamb & gimmers balance				-0.55	1	1	1														≤	0
5. Sell for gimmering						-1			-1									1			≤	0
6. Buy gimmers												0.28						-0.98			≤	0
7. Lowground ewes balance												-1.44	1	1							≤	0
8. Draft blackface from Lowground													-1		1						≤	0
9. Draft blackface Ewes															-1.35	1	1				≤	0
10. Store Lambs			-1				-1			-1						-1			1	1	≤	0

Lambs produced by hill breeding ewes that are not intended for replacements can either be sold to upland farms for breeding, which need to replace a quarter of their breeding ewes, or to other units specialised in fattening (balance rows 1, 2, 10). Lambs produced by other hill ewes that are not intended for replacements can be either finished or sold for gimmering (balance rows 4, 5, 10). Upland farms either sell their lambs for gimmering, either store or finish them (balance rows 3, 5, 10). Lowground breeding ewes produce lambs for drafting or finishing (balance rows 7, 8). Drafted ewes produce lambs for storing or finishing (balance rows 9, 10). Gimmering units sell their lambs to lowground farms (balance row 6). Finally, stored lambs are sold for finishing either in winter or in autumn (balance row 10).

4.3.3. Mathematical Structure and Assumptions

Input and output prices are exogenous. The endogenous variables of the model are the agricultural activities. The solution is the mix of the activities that optimises the objective function. In its standard form the model is formulated algebraically as follows:

Find the activity vector $\mathbf{x} = (x_1, x_2, \dots, x_j)$

that maximises $$TGM^s = \sum_{j=1}^{J} (gm_j + sub_j) \cdot x_j - \sum_{j=1}^{J} C_j(x_j) \qquad (1a)$$

Subject to $$\sum_{j=1}^{J} a_{kj} x_j \le b_k, \quad all\ k = 1\ to\ K \qquad (1b)$$

$$x_j \ge 0, \quad all\ j = 1\ to\ J \qquad (1c)$$

In accordance to the PMP rationale that variable marginal costs include accounting costs, c as well as the “hidden” or opportunity costs incorporated in the shadow values of calibrated activities, $\lambda^{cal(LP)}$, the objective function (4.1) to be maximized is total gross margin (TGM^s)[1] which includes shadow values.

[1] A total gross margin function that uses the “true” economic prices of the activities *i.e.* captures all the shadow variables involved in deciding the optimum level of each activity, not merely those for which market prices exist. In the present PMP model these shadow variables are represented by the opportunity cost of the activities.

The gross margin coefficient, gm_j, for an output activity j is defined as the difference between its revenue and the sum of all of its variable production costs that are not represented by the input activities in the model. Gross margins, gm_j, may be supplemented by different kinds of subsidies, sub_j. The total variable costs, $C_j(x_j)$, of each activity, x_j, include accounting variable costs, c_j, and opportunity costs (shadow values), $\lambda^{cal(LP)}$. These are represented by a nonlinear function of all or some of the activities in the model. The specification of the nonlinear function is explained later on. In the constraint equations (1b) a_{kj} is the coefficient that expresses the amount of input k required per unit of output activity j or provided per unit of input activity j, and k denotes the index for constraints. The constraints, b_k, comprise the total available land, which is assumed to be the only limiting factor of production, the policy constraints, the balances among livestock activities, the balances between forage and livestock activities, and the balances between other intermediate inputs and factors and output activities.

The model assumes that there are cross effects only between the cropping activities, only between the cattle activities and only between sheep activities but not between activities of these different groups. Additionally, it assumes that the opportunity costs of cropping activities reflect also the differences across land classes and climatic conditions. This means that the opportunity cost of a crop increases as more land of worse quality is allocated to that crop. Treating Scottish farming systems as a single-macro farm implies a high degree of labour and capital mobility between farms. It is assumed that it is a closed market with respect to livestock purchases but it is possible to purchase feed from outside the country and to transfer forage within different farms of different regions.

Following the PMP paradigm presented in Chapter 3, a specification of the total variable cost function, $C_j(x_j)$, being the integral of the variable marginal costs ($\lambda_j^{cal(LP)} + c_j$) with respect to x_j^R over the interval (0, x_j^R) reads

$$C_j(x_j) = \int (\lambda_j^{cal(LP)} + c_j)' dx_j \qquad (2)$$

The choice of different functional specifications for the marginal cost corresponds to different but admissible specifications of the total variable function. In the present model the familiar quadratic total variable cost function is used which is the result of postulating that

$$\int_0^{x_j^R} (\lambda_j^{cal(LP)} + c_j)'\partial x_j = \frac{1}{2} q_{jj} x_j^2 + \sum_{i \neq j} q_{ji} \cdot x_j \cdot x_i \quad \forall \quad j$$

$$\lambda_j^{cal(LP)} + c_j = \frac{\partial(\frac{1}{2} q_{jj} x_j^2 + \sum_{i \neq j} q_{ji} \cdot x_j \cdot x_i)}{\partial x_j} = q_{jj} x_j + \sum_{i \neq j} q_{ji} x_i \quad \forall \quad j$$

or in matrix notation[2]:

$$(\boldsymbol{\lambda}^{\mathrm{cal(LP)}} + \mathbf{c}) = \mathbf{Q}\ \mathbf{x},$$

Therefore the total variable cost function now reads:

$$\boldsymbol{C}(\mathbf{x}) = \frac{1}{2} \mathbf{x}'\mathbf{Q}\mathbf{x} \qquad (3)$$

with $\boldsymbol{Q}$ a symmetric positive definite matrix of parameters associated with the quadratic term. In this case we have to distinguish the diagonal elements of $\boldsymbol{Q}$, q_{jj}, which express the change in marginal cost of activity j with respect to the level of activity j from the off-diagonal elements, q_{ji}, which express the change in marginal cost of activity i with respect to the level of activity j. The reconstruction of the $\boldsymbol{Q}$ matrix can be achieved by using either the standard PMP calibration method, the ME recovery method or the GME estimation method as shown in the next section. All three approaches use a first step where the internal shadow values, $\lambda^{cal(LP)}$, are derived from an LP model. Then the validation and comparison of the different methods of parameter calibration, recovery and estimation are presented in the last section.

Under various changes in the prices, subsidy rates, variable costs, and constraint levels the PMP model finds the levels of the activities that maximise the total gross margins.

4.4. Estimation of the Q Matrix

Linear cost functions are not replaced for all activities with equivalent quadratic specifications Only those that theory or data suggest are best modelled as nonlinear

---

[2] For example, when $j = 3$ we have:

$$\begin{array}{lll} \lambda_1^{cal} + c_1 = q_{11}x_1 + q_{12}x_2 + q_{13}x_3 & & \lambda_1^{cal} + c_1 = \mathbf{q}_1 \cdot \boldsymbol{x} \\ \lambda_2^{cal} + c_2 = q_{22}x_2 + q_{21}x_1 + q_{23}x_3 & \Rightarrow & \lambda_2^{cal} + c_2 = \mathbf{q}_2 \cdot \boldsymbol{x} \ \Rightarrow \ (\boldsymbol{\lambda}^{cal} + \boldsymbol{c})_{(3\mathrm{x}1)} = \boldsymbol{Q}_{(3\mathrm{x}3)}\, \boldsymbol{x}_{(3\mathrm{x}1)} \\ \lambda_3^{cal} + c_3 = q_{33}x_3 + q_{31}x_1 + q_{32}x_2 & & \lambda_3^{cal} + c_3 = \mathbf{q}_3 \cdot \boldsymbol{x} \end{array}$$

are replaced. Activities used as intermediate inputs (that are not marketed in the model) such as pastures and fodder crops grown to be fed directly to livestock are not calibrated but rather their levels are determined by the demand for those activities from other PMP-calibrated activities. The amount of forage each animal needs is defined as a fixed input requirement since linear production technology is assumed. The cash crops are calibrated to the land use, while the cattle and sheep activities are calibrated to the observed numbers. The area of grassland is recorded in the agricultural census. However, information on which activities are actually carried out and to what extent is not available. For example, grassland kept for grazing, hay, silage is not distinguished.

The **Q** matrix that results from the selection process needs to be recovered. In Chapter 3 three approaches were presented. One is the standard PMP method which specifies the elements of **Q** in such a way as the model calibrates to a set of base year observations. The ME and GME methods which use one and multiple observation sets respectively to estimate the elements of **Q** are the other two. The remainder of this empirical chapter first presents and then validates the results for **Q** from each of these methods before concluding with a short discussion.

4.4.1. Standard PMP method (calibration)

As already stated in Chapter 3, the amount of information available for reconstructing the **Q** matrix is given by the vector ($\boldsymbol{\lambda}^{\mathrm{cal}} + \mathbf{c}$) and the vector of observed output levels, $\mathbf{x}^R$. Setting all off-diagonal elements of **Q** to zero, the $\boldsymbol{N}$ diagonal elements of **Q**, q_{jj}, can then be calculated (calibrated) as

$$q_{jj} = \frac{\lambda_j^{cal} + c_j}{x_j^R}\text{, for all } i = 1, 2, \ldots, \boldsymbol{N} \qquad (4)$$

For the arable activities, $\boldsymbol{x}_{\boldsymbol{\alpha}}$, the observed levels of the subvector, $\boldsymbol{x}_{\boldsymbol{\alpha}}^{\boldsymbol{R}}$, are taken directly from the agricultural census data. However, for the livestock activities as specified in the model the observed levels are not available in the census dataset. Only total numbers of the dairy cows, beef cattle, calves, ewes and lambs are known for the whole country. Hence, structural relationships between the various livestock activities (such as proportions of animals kept for breeding for storing and for finishing) were used in order to derive estimations for the actual levels of each activity.

Tables 4.4, 4.5 and 4.6 present the calibrated diagonal parameters (the off-diagonal elements were assumed zero) of the quadratic cost function for arable, cattle and sheep activities respectively. These are derived from expression (4) for the base year 2004. When the objective function of model (1a)-(1c) with the PMP-calibrated parameters is maximized, the model almost exactly replicates the actual shares of crop and livestock activities existing under the base year conditions (determined by yields, commodity prices, variable costs and policy measures).

Nonetheless, as explained in Chapter 3, there are two significant practical problems with the diagonal specification in that it assumes that there are no substitution or complementarity cost effects between activities in the same district or farm, apart from the effect on the total land constraint. Clearly, the existence of rotations in crop production and linkages between livestock activities (breeding-storing-wintering-finishing) implies that the model should take into account the interdependencies among production activities, and include relevant parameters into the objective function.

Table 4.4. q_{ii} parameters for the Quadratic Cost Function of cropping activities calibrated by the standard PMP method

Crop	Setaside	wWheat	wBarley	sBarley	wOats	sOats	wOSR	sOSR	Triticale
Setaside	**0.002633**								
wWheat		**0.005203**							
sBarley			**0.010985**						
sBarley				**0.001958**					
wOats					**0.107642**				
sOats						**0.030360**			
wOSR							**0.013403**		
sOSR								**0.072586**	
Triticale									**0.343750**

Base year: 2004

Table 4.5. q_{ii} parameters for the Quadratic Cost Function of cattle activities calculated by the standard PMP method

Cattle activity	1	2	3	4	5	6	7	8	9	10	11
1. Dairy Cows (Spring –Summer calving)	**5.70×10^{-3}**										
2. Dairy Cows (Autumn calving)		**9.76×10^{-3}**									
3. Hill Suckler Cows			**2.73×10^{-3}**								
4. Upland/Lowground Cattle (silage / Fb-Jn)				**2.18×10^{-3}**							
5. Upland/Lowground Cattle (silage / Ag-Oc)					**3.75×10^{-3}**						
6. Upland/Lowground Cattle (straw / Fb-Jn)						**1.23×10^{-3}**					
7. Upland/Lowground Cattle (straw / Ag-Oc)							**3.39×10^{-3}**				
8. Overwintering (Spring Calves)								**0.16×10^{-3}**			
9. Intensive Finishing (Spring Calves at 12 m)									**2.19×10^{-3}**		
10. Finishing (Autumn Calves)										**0.45×10^{-3}**	
11. Finishing (Spring Calves at 18-24 m)											**0.49×10^{-3}**

Base year: 2004

Table 4.6. q_{ii} parameters for the Quadratic Cost Function of sheep activities calculated by the standard PMP method

Sheep activity	1	2	3	4	5	6	7	8
1. Hill breeding ewes (store lamb production)	**0.000018**							
2. Hill breeding ewes (finish lamb and gimmer production)		**0.000057**						
3. Upland breeding ewes			**0.000051**					
4. Lowground breeding ewes (early finished lamb)				**0.000025**				
5. Draft hill ewes (finish and store lambs)					**0.000036**			
6. Gimmering						**0.000022**		
7. Winter finishing lambs							**0.000049**	
8. Autumn finishing lambs								**0.000002**

Base year: 2004

4.4.2. Maximum Entropy (estimation with one year observations)

The assumption of diagonal **Q** is unrealistic but to calibrate the full matrix of coefficients means that the number of unknown parameters of the full matrix, [$\boldsymbol{N}(\boldsymbol{N}+1)/2$], is larger than the number of observations, $\boldsymbol{N}$ and the problem is ill-posed. If independency is assumed between the activities of different groups (without imposing significant error), the number of unknown parameters is reduced dramatically. Therefore, three submatrices, $\mathbf{Q}_a$, $\mathbf{Q}_c$ and $\mathbf{Q}_s$, one for each activity category (arable, cattle, sheep) can be recovered instead of the whole **Q**. In Chapter 3 it was demonstrated how the ME principle provides the analytical framework for solving ill-posed problems like this. Applying the ME for each activity group yields three of the following problems as formulated in Paris and Howitt (1998):

$$\max H(\mathbf{p}^{L}, \mathbf{p}^{D}) = - \sum_{jj',k} \mathbf{p}^{L} \log \mathbf{p}^{L} - \sum_{jj,k} \mathbf{p}^{D} \log \mathbf{p}^{D} \tag{5a}$$

subject to

$$\begin{aligned}(\boldsymbol{\lambda}^{cal(LP)} + \mathbf{c}) = \mathbf{mc} &= \mathbf{Qx}^{R} \\ &= \mathbf{LDL'x}^{R} \\ &= (\mathbf{Z^{L}p^{L}})(\mathbf{Z^{D}p^{D}})(\mathbf{Z^{L}p^{L}})' \, \mathbf{x}^{R}\end{aligned} \tag{5b}$$

$$\sum_{k} p^{L}_{(j,j',k)} = 1 \tag{5c}$$

$$\sum_{k} p^{D}_{(j,j',k)} = 1 \tag{5d}$$

The ME problem consists in recovering the $\boldsymbol{p}^{\boldsymbol{L}}$ and $\boldsymbol{p}^{\boldsymbol{D}}$ probability matrices. Given that there are say, J activities of each type (crops, cattle, sheep) and therefore (J x J) elements of each $\mathbf{Q}$ matrix and given that each element $q_{jj'}$ is specified with K support values, the $\mathbf{Z}^{\boldsymbol{L}}$ and $\mathbf{Z}^{\boldsymbol{D}}$ matrices are specified according to Paris and Howitt (1998) as follows:

$$\text{for } j = j' \qquad Z^{D}_{(j,j',k)} = \frac{mc_j}{x^{R}_{j}} w^{D}_{k} \quad k = 1, 2, \ldots, K \tag{6}$$

$$\text{for } j \neq j' \qquad Z^{D}_{(j,j',k)} = 0 \tag{7}$$

$$\text{for } j > j' \qquad Z^{L}_{(j,j',k)} = \frac{mc_j}{x^{R}_{j}} w^{L}_{k} \quad k = 1, 2, \ldots, K \tag{8}$$

$$\text{for } j = j' \qquad Z^{L}_{(j,j',k)} = 1 \tag{9}$$

where $\mathbf{w^{D}}$ and $\mathbf{w^{L}}$ are (K x 1) vectors of suitable support weights[3]. The quantity x^{R}_{j} is the realised output level of the j^{th} activity. The quantity mc_j is the j^{th} marginal cost which can be measured in the LP (first) stage of the PMP procedure or derived from some other source of information. The base year is 2003. The recovered parameters of the **Q** matrix (and consequently the behaviour of the model) depend on the end-point values of the support set (Golan *et al*, 1996). In this study, the selection of the support end points was made on the basis of model performance. After testing several alternative sets of support weights, the final set $\mathbf{w^{D}}$ of positive (to ensure that the

[3] For the purpose of this application it suffices to consider common sets of support values for the unknown parameters; a practice followed by others too for example, Léon et al., 1999 (pp 428).

resulting **Q** matrix be positive semi-definite) weights for the diagonal elements of the **D** matrix is (0, 1, 2, 3, 4), and the final set $\mathbf{w}^{\mathbf{L}}$ of weights for the off-diagonal elements of the **L** matrix is (-1, -0.5, 0, 0.5, 1).

Problem (5a)-(5d) was solved by using a programme developed in C++ programming language by the author. The recovered matrices $\mathbf{Q}_a$, $\mathbf{Q}_c$, $\mathbf{Q}_s$ are presented in Tables 4.7, 4.8 and 4.9 respectively. It is worth pointing out that the off-diagonal parameter estimates of **Q** matrices are found equal to zero, the centre of the set of support weights $\mathbf{w}^{\mathbf{L}}$. This happens because $H(\mathbf{p}^{L}, \mathbf{p}^{D})$ reaches a maximum when there is maximum uncertainty in relation to the values of the parameters $Z^{L}_{(j,\ j',k)}$. In other words, for all parameters the probability of any support point is $p^{L}_{(j,\ j',k)} = 1/k$, which gives an estimation of zero. Another point is that, lacking any additional information, ME gives the same results as PMP. Of course when more reliable information becomes available it can be incorporated in the support values, whereas with the standard PMP approach this is not possible. Therefore, ME does not necessarily need an auxiliary LP model when there are other sources of information regarding the possible range of values that marginal costs, mc_j, can take.

When the support values depend entirely on an auxiliary LP, as is the case in this study, they will be different for different base years. It is clear that changes in the production activities between years will generate different costs, and consequently different dual values, λ_j^{cal}. Because these are incorporated in marginal costs, mc_j, used for the estimation of **Q**, for different base years, different **Q** matrices will be estimated. This is a potential problem because the model's behaviour is ultimately determined by this rather arbitrary specification.

Table 4.7. Recovered Qa (symmetric) matrix for the quadratic Cost Function of cropping activities recovered with the ME method

Crop	Setaside	wWheat	sBarley	sBarley	wOats	sOats	WOSR	sOSR	Triticale
Setaside	**0.002012**	0	0	0	0	0	0	0	0
wWheat		**0.005450**	0	0	0	0	0	0	0
sBarley			**0.010688**	0	0	0	0	0	0
sBarley				**0.001831**	0	0	0	0	0
wOats					**0.10818**	0	0	0	0
sOats						**0.028828**	0	0	0
wOSR							**0.013746**	0	0
sOSR								**0.064583**	0
Triticale									**0.35374**

Base year: 2003

Table 4.8. Recovered Qc (symmetric) matrix for the Quadratic Cost Function of cattle activities calculated by the ME method

Cattle activity	1	2	3	4	5	6	7	8	9	10	11
1. Dairy Cows (Spring –Summer calving)	**0.00502**	0	0	0	0	0	0	0	0	0	0
2. Dairy Cows (Autumn calving)		**0.00872**	0	0	0	0	0	0	0	0	0
3. Hill Suckler Cows			**0.00176**	0	0	0	0	0	0	0	0
4. Upland/Lowground Cattle (silage / Fb-Jn)				**0.00128**	0	0	0	0	0	0	0
5. Upland/Lowground Cattle (silage / Ag-Oc)					**0.00214**	0	0	0	0	0	0
6. Upland/Lowground Cattle (straw / Fb-Jn)						**0.00166**	0	0	0	0	0
7. Upland/Lowground Cattle (straw / Ag-Oc)							**0.00376**	0	0	0	0
8. Overwintering (Spring Calves)								**0.0001**	0	0	0
9. Intensive Finishing (Spring Calves at 12 m)									**0.00234**	0	0
10. Finishing (Autumn Calves)										**0.0006**	0
11. Finishing (Spring Calves at 18-24 m)											**0.00048**

Base year: 2003

Table 4.9. Recovered Qs (symmetric) matrix for the Quadratic Cost Function of Sheep activities calculated by the ME method

Sheep activity	1	2	3	4	5	6	7	8
1. Hill breeding ewes (store lamb production)	**$2.4x10^{-5}$**	0	0	0	0	0	0	0
2. Hill breeding ewes (finish lamb and gimmer production)		**$8.08x10^{-6}$**	0	0	0	0	0	0
3. Upland breeding ewes			**$6.74x10^{-6}$**	0	0	0	0	0
4. Lowground breeding ewes (early finished lamb)				**$2.78x10^{-6}$**	0	0	0	0
5. Draft hill ewes (finish and store lambs)					**$7.68x10^{-6}$**	0	0	0
6. Gimmering						**$5.64x10^{-6}$**	0	0
7. Winter finishing lambs							**$4.02x10^{-6}$**	0
8. Autumn finishing lambs								**$0.49x10^{-6}$**

Base year: 2003

4.4.3. Generalised Maximum Entropy (estimation with multiple observations)

As already recognised by Paris and Howitt (1998) and by Heckelei and Britz (2000) one way to reduce the effect of the support end points, and effectively the arbitrary behaviour of the model, is to use more observations. As more observations become available, more reliable estimates of the parameters of the total variable cost function can be achieved by using the GME approach. The problem is then specified as follows:

$$\max H(\mathbf{p}^{L}, \mathbf{p}^{D}, \mathbf{v}) = -\sum_{j,j',k} \mathbf{p}^{L} \log \mathbf{p}^{L} - \sum_{j,j,k} \mathbf{p}^{D} \log \mathbf{p}^{D} - \sum_{t} \mathbf{v}_{t} \log \mathbf{v}_{t} \qquad (10a)$$

subject to

$$\boldsymbol{mc_t} = \mathbf{Q}(\mathbf{x}^{R}{}_{t} + \mathbf{e})$$

$$= \mathbf{LDL'}(\mathbf{x}^{R}{}_{t} + \mathbf{e})$$

$$= (\mathbf{Z^L p^L})(\mathbf{Z^D p^D})(\mathbf{Z^L p^L})' \, (\mathbf{x}^{R}{}_{t} + \mathbf{V}\mathbf{w}_{t}) \; \forall \; t=1, 2,\ldots,5 \qquad (10b)$$

$$\sum_{k} \mathbf{p}_{k}^{L} = 1 \qquad (10c)$$

$$\sum_{k} \mathbf{p}_{(j,j',k)}^{D} = 1$$

(10d)

$$\sum_{k} v_{t,k} = 1 \qquad (10e)$$

where t is the year, and the last three constraints express the adding-up condition of probabilities. In this application, observations from three years, 2000, 2002 and 2004 were used. For the range of the error support set, Golan *et al* (1996, pp. 139-140) suggest the 3σ rule, where σ is the standard deviation of the observations. However, solving the above problem with a (-3σ, 3σ) range did not result in a feasible solution for cattle and sheep activities. This was due to the fact that big differences in **mc** from year to year generated large errors. Thus, the range of the error support set that was finally used for recovering matrices $\mathbf{Q}_{c}$, and $\mathbf{Q}_{s}$ was ±20 standard deviations[4].

Problem (10a)-(10d) was also solved by using a programme developed in C++ programming language by the author. The recovered matrices $\mathbf{Q}_{a}$, $\mathbf{Q}_{c}$, $\mathbf{Q}_{s}$ are presented in Tables 4.10, 4.11 and 4.12 respectively. In this case, some of the off-diagonal estimates are non zeros due to the effect of the multiple observations. In terms of how the model's variables (production activities) are interrelated, the data "force" the estimation procedure to reconstruct **Q** matrices that incorporate cross-effects between the activities. But even so, the **Q** matrices may not be able to "explain" entirely the behaviour of the "true" model which has produced the multiple

[4] There is a controversy about the impact of the support bounds on parameter estimates. In their seminal book Golan *et al*. (1996) suggest that widening the parameter and error supports has little effect on the estimates. This conclusion has been reiterated by the study of Lence and Miller (1998). However, Paris and Caputo (2001) in their working paper on the sensitivity of the GME estimates to support bounds present empirical evidence which argues that the impact of variations of parameter and error support bounds is unpredictable.

observations. In order for the estimated **Q** matrices to satisfy the data constraints and thus a feasible solution to exist, disturbances are needed (vector **e** in equation 10b).

Table 4.10. Estimated Q (symmetric) matrix for the Quadratic Cost Function of cropping activities recovered with the GME method

Crop	Setaside	wWheat	wBarley	sBarley	wOats	sOats	wOSR	sOSR	Triticale
Setaside	**0.002441**	-0.000005	-0.000006	-0.000013	0.000000	-0.000003	-0.000001	0.000000	-0.000001
wWheat		**0.005493**	-0.000001	-0.000010	0.000000	0.000001	0.000012	0.000003	-0.000002
sBarley			**0.011281**	-0.000016	-0.000004	0.000000	0.000075	0.000036	-0.000009
sBarley				**0.002121**	0.000000	-0.000001	0.000004	0.000002	-0.000002
wOats					**0.117429**	-0.000095	0.000159	-0.000005	-0.000069
sOats						**0.033577**	0.000081	0.000102	-0.000019
wOSR							**0.014529**	0.000022	-0.000005
sOSR								**0.065450**	0.000064
Triticale									**0.369228**

Observation years: 2000-2002-2004

Table 4.11. Estimated Q (symmetric) matrix for the Quadratic Cost Function of cattle activities recovered with the GME method

Cattle activity	1	2	3	4	5	6	7	8	9	10	11
1. Dairy Cows (Spring –Summer calving)	**0.008674**	-0.00010	-0.00008	-0.00018	-0.00007	-0.00018	-0.00010	-0.00030	-0.00007	0.00027	-0.00039
2. Dairy Cows (Autumn calving)		**0.014026**	-0.00009	-0.00023	-0.00008	-0.00024	-0.00013	-0.00034	-0.00010	0.00050	-0.00050
3. Hill Suckler Cows			**0.004290**	-0.00004	-0.00003	-0.00003	-0.00003	-0.00017	-0.00001	-0.00023	-0.00013
4. Upland/Lowground Cattle (silage / Fb-Jn)				**0.003132**	-0.00003	-0.00003	-0.00002	-0.00015	-0.00001	-0.00017	-0.00011
5. Upland/Lowground Cattle (silage / Ag-Oc)					**0.006696**	-0.00007	-0.00005	-0.00030	-0.00002	-0.00035	-0.00022
6. Upland/Lowground Cattle (straw / Fb-Jn)						**0.001794**	-0.00001	-0.00005	0.00000	-0.00009	-0.00003
7. Upland/Lowground Cattle (straw / Ag-Oc)							**0.004076**	-0.00014	-0.00001	-0.00017	-0.00010
8. Overwintering (Spring Calves)								**0.000598**	0.00001	0.00003	0.00005
9. Intensive Finishing (Spring Calves at 12 m)									**0.003954**	-0.00002	-0.00004
10. Finishing (Autumn Calves)										**0.000648**	0.00001
11. Finishing (Spring Calves at 18-24 m)											**0.001174**

Observation years: 2000-2002-2004

Table 4.12. Estimated Q (symmetric) matrix for the Quadratic Cost Function of Sheep activities recovered with the GME method

Sheep activity	1	2	3	4	5	6	7	**8**
1. Hill breeding ewes (store lamb production)	**0.000047**	-0.000002	-0.000004	-0.000008	-0.000003	-0.000003	-0.000003	-0.000005
2. Hill breeding ewes (finish lamb and gimmer production)		**0.000115**	0.000002	-0.000013	-0.000009	0.000006	0.000003	0.000011
3. Upland breeding ewes			**0.000104**	-0.000004	0.000005	-0.000010	-0.000014	-0.000014
4. Lowground breeding ewes (early finished lamb)				**0.000035**	-0.000001	0.000000	0.000000	0.000002
5. Draft hill ewes (finish and store lambs)					**0.000093**	-0.000004	-0.000005	-0.000001
6. Gimmering						**0.000062**	0.000001	-0.000002
7. Winter finishing lambs							**0.000092**	-0.000009
8. Autumn finishing lambs								**0.000016**

Observation years: 2000-2002-2004

4.5. Validation of Models

If the model is to capture effectively the process of farmers' production choices, it must be demonstrably reliable. This means that the model has to be validated. McCarl and Apland (1986) identify two validation approaches: validation by construct and validation by results. Validation by construct asserts the model was built properly therefore it is valid. The model is justified valid by construct when *i*) the right procedures were used by the model builder; *ii*) the results do not contradict a priori perceptions of reality and, *iii*) constraints were imposed which restrict the model to replicate an observed outcome. In all cases however validation of a particular model is assumed, not tested. Validation by results systematically compares the model output against real world observations. McCarl and Apland (1986) suggest that at least five sequential validation experiments should be performed: a feasibility experiment, a quantity experiment, a price experiment, a prediction experiment, and a change experiment. Model Validation is fundamentally objective nonetheless it reveals model strengths and weaknesses which are valuable to users (McCarl and Apland, 1986).

The prediction experiment is the most commonly used experiment for validating LP models by results. This particular experiment is performed to see whether the model can replicate farmer's behaviour (reflected in their production decisions) that has been observed. Such observations on past agricultural land use (land cover but not land management) were derived from the June agricultural census.

Three different models resulted from the three different estimation procedures of the **Q** matrix. The first derived from the standard PMP approach with base year 2004, the second from the ME approach with base year 2004, and the third from the GME. The models are solved with **Solver DLL**, an optimisation software package developed by Frontline Systems, Inc. The models were run five times one for each year from 2000 to 2004 using actual prices and policies in place, and the simulated and observed land use patterns were compared.

The performance of models can be evaluated by a number of ways. One way is to show graphically the ability of the three models to predict past observations. Figures 4.5, 4.6 and 4.7 present this information for arable, cattle and sheep activities respectively. The bars represent percentage deviations ((observed – predicted) / observed) between the predicted and the observed levels for each activity. Bars corresponding to different years are clustered together for each activity. The quality of the predictive ability of a model can also be evaluated using various measures. For example, Heckelei and Wolff (2003) use *absolute bias* (ABIAS = absolute value of the difference between average estimate across sample and true value of the parameter) and *root mean square error* (RMSE = square root of the mean squared distance between estimates and true parameter). Moxey *et al*. (1995) used the *percentage absolute deviation* (PAD) goodness-of-fit measure which is defined as the average absolute deviation between the predicted and the observed levels (areas for crops and heads for livestock) divided by the average actual value (see also Norton & Schiefer, 1980). PAD values were used here to evaluate and compare the performance of the three models. The results are summarised in Table 4.13.

The arbitrary behaviour of the PMP and ME derived models, discussed earlier, is manifested in the differences in PAD values between those derived from 2003 observations and those derived from 2004 observations. In particular, the "2004-based" model appears to fit better the set of observations possibly because 2004 happens to be a better representative of all years in the set than 2003.

One striking observation is that predictions for cattle and to a higher extent for sheep are significantly worse than those for crops. One possible explanation relates to the quality of the data used as observations. As already mentioned in section 4.4, observations for the livestock activities, as they are defined in the model, were not readily available from the agricultural Census; instead they had to be derived indirectly using assumption about their interlinkages. One the other hand, data about marginal costs taken from the Farm Management Handbook can be regarded as more reliable. Both types of data are incorporated in the model in such a way that it is assumed that they are correlated namely, activity levels depend on marginal costs. If livestock activity data, which were derived indirectly, are poor approximation of the data that the marginal costs used would have generated, then they may have biased the estimation procedure.

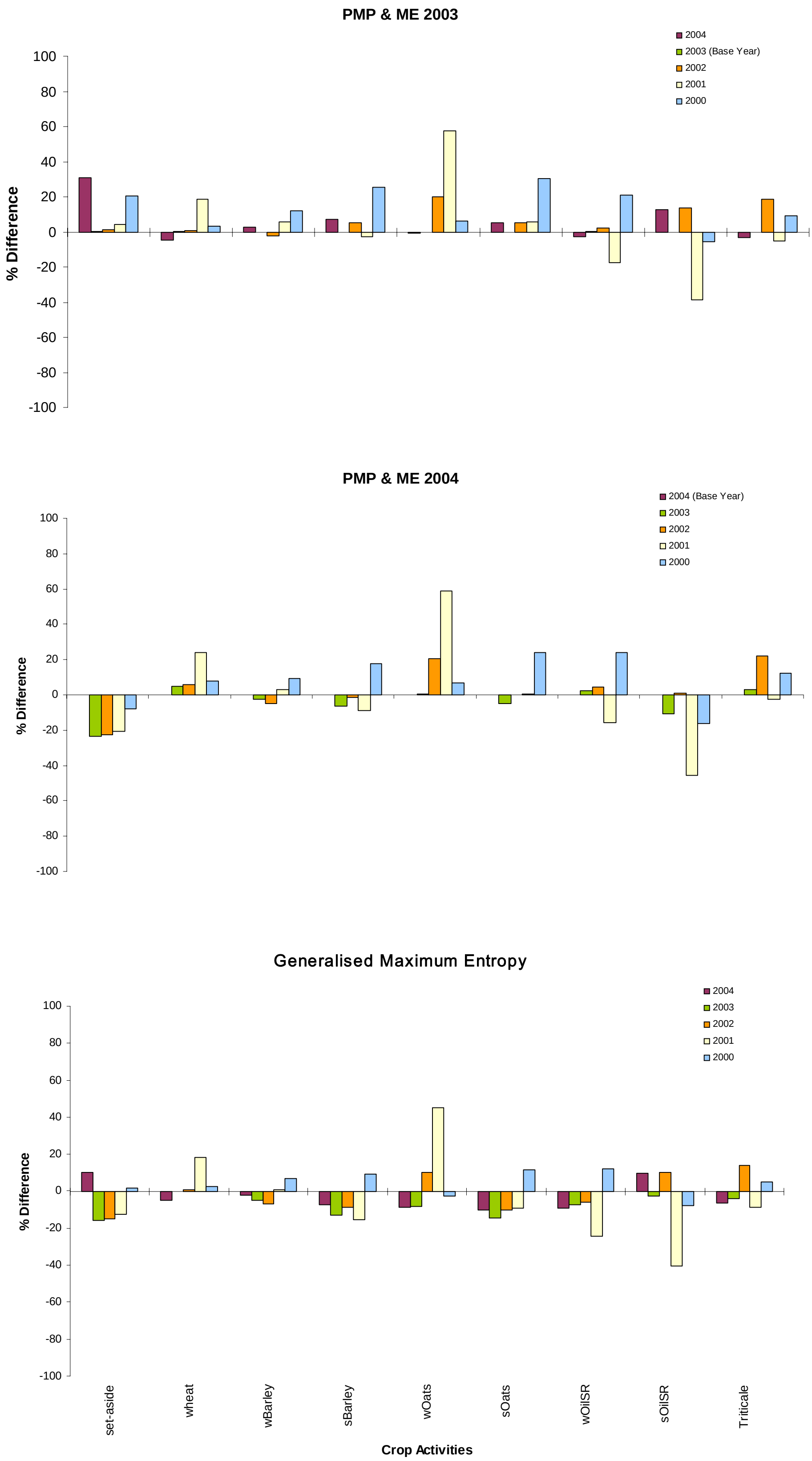

Figure 4.5. Percentage deviations between the observed levels of arable activities and their levels as predicted by the four different models

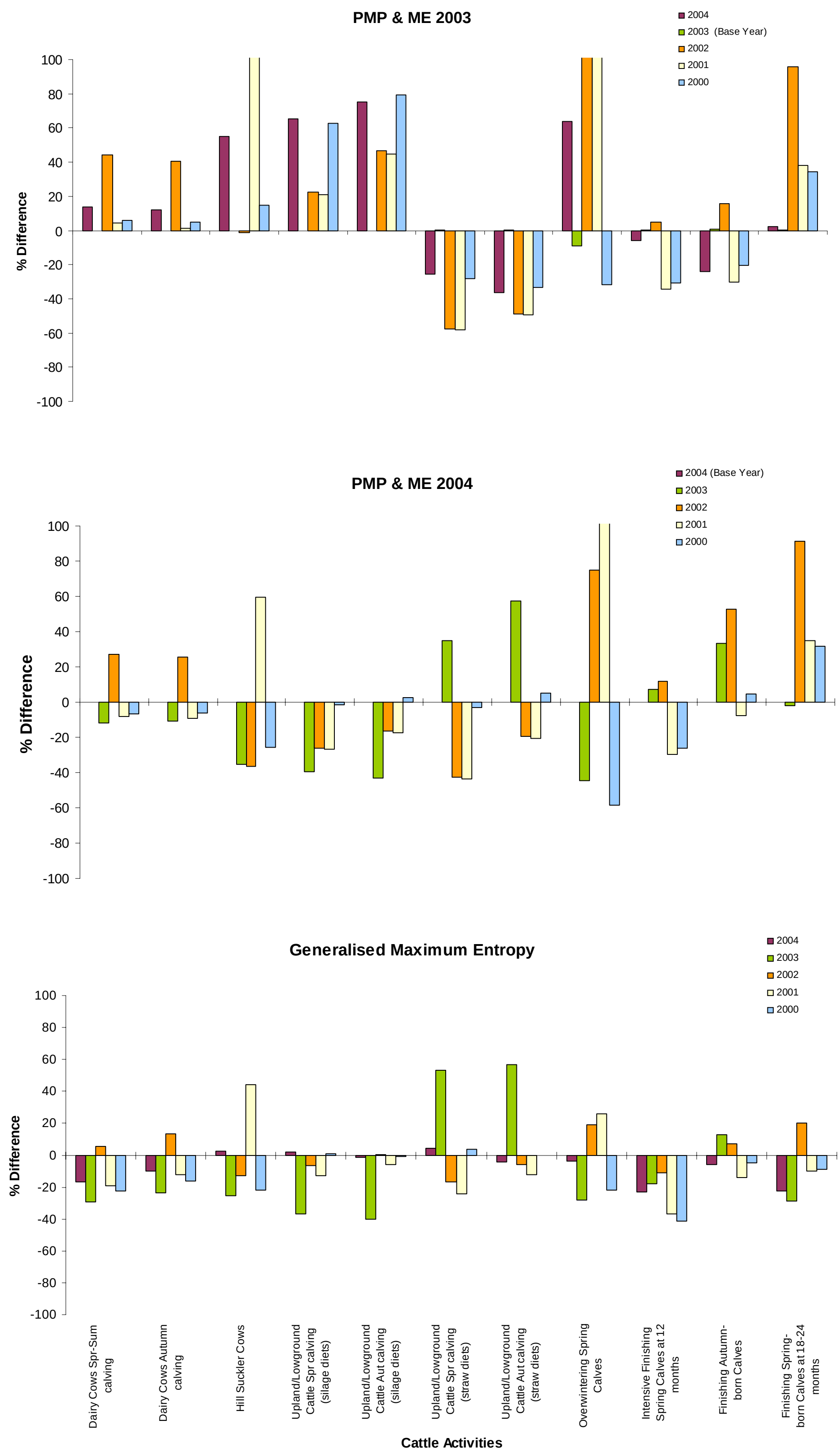

Figure 4.6. Percentage deviations between the observed levels of cattle activities and their levels as predicted by the four different models

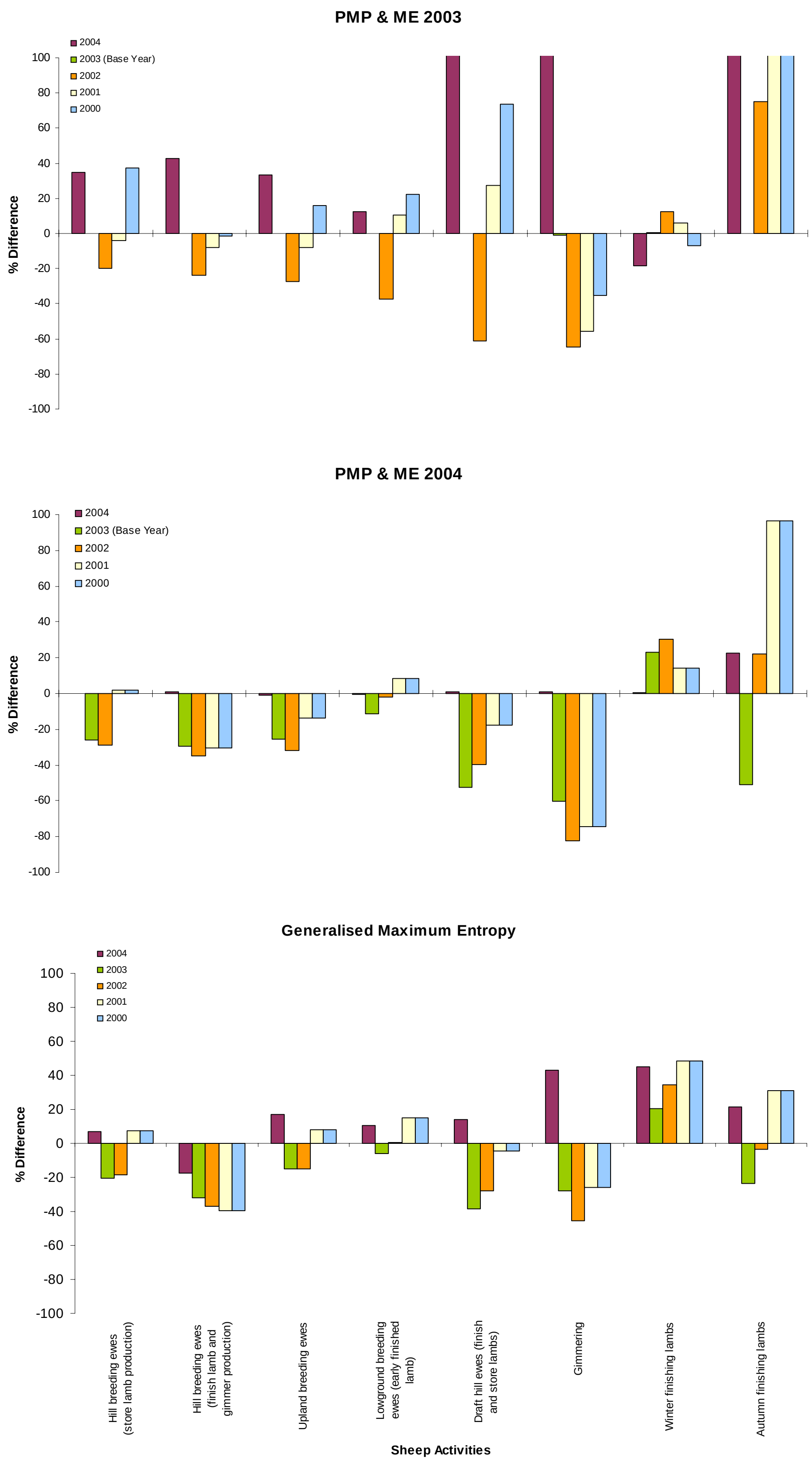

Figure 4.7. Percentage deviations between the observed levels of sheep activities and their levels as predicted by the four different models

The standard PMP and the ME procedures generated the same **Q** matrices. PMP and ME are both based on one year observations, which allows the models derived from them to calibrate to this year. This means that the percentage deviations of both for their base year are very close to zero. This is also illustrated by the almost flat bars for all activities. However, there is no mechanism to take under consideration data from the other years. This explains why PAD values are higher when the base year is not included for their computation (columns 3&5 in Table 4.13). This however, is not shown when comparing the **Q** matrices derived from the standard PMP and the ME method simply because their base years are different. This choice was made intentionally in order to demonstrate two things: first, the subjectivity of the method expressed in the choice of support values and second, that there are more than one **Q** specification with which the model calibrates.

Table 4.13. PAD values for the three models

	PMP/ME (2003)		PMP/ME (2004)		GME
	Base year not included	Base year included	Base year not included	Base year included	
crops	10.51	8.41	10.81	9.97	10.10
cattle	36.20	44.99	29.44	23.55	16.88
sheep	50.83	40.72	37.61	30.77	26.52
mean	**32.52**	31.37	**25.95**	21.43	**17.83**
stdev	20.41	20.00	13.74	10.56	8.25

Note: Different PAD values derived from the ME estimated model and the standard PMP calibrated model do not mean than the ME is a superior method.
Different base years were used and the differences are due to the fact that the year used for the ME happened to be more “representative” than that used for the standard PMP.

Incorporating more observations in the estimation procedure can improve the estimation of the model’s parameters, given that the additional data are reliable. This can be done by applying the GME method. The mean PAD value and standard deviation are lowered almost by 20% making the model’s predictions significantly more accurate. This finding agrees with the expectation that more observations reduce the variation in the observations that is not accounted for by the estimation of the systematic component of the model (matrix **Q** in equation 10b). What appears puzzling, however, is the slightly lower PAD value of the GME method for the arable crop activities compared to PMP/ME when the base year is included. This can be due to the combination of two factors. The first one is that single observation *mc* values are already reasonably good estimates (in a sense that they produce low residuals). The second is that GME imposes a priori a range of values for each of these errors to relax the constraints given by equation (10b). If their assumed range is larger compared to the error values the PMP and ME calibrations produce, the GME-derived errors will be higher too. Consequently, in situation like this GME can add rather than reduce noise in the estimation.

In order to evaluate the effect of model calibration method on the performance of the calibration an analysis of variance (ANOVA) was carried out on the PAD values. For each calibration method, the available data for the analysis were the PAD values for each of the five years and for each of the three main types of activity (cattle, crops and sheep) within the standard SASEM model. The full interaction model for the factors calibration method and activity was fitted, using years as replicates.

The summary table from the ANOVA is presented in Table 4.14. It was found that the mean PAD value differed between activities (F_{pr} = 0.001, (d.f. = 2, 32)) but not calibration methods ((F_{pr} = 0.161, (d.f. = 2, 32)), nor was there any evidence of an significant interaction between activity and calibration method (F_{pr} = 0.541, (d.f. = 4, 32)). Irrespective of calibration method the mean PAD for crops (10.7%) was lower than for either cattle (35.1%) or sheep (32.7) (l.s.d. at p = 0.05 = 13.61).

Table 4.14. Analysis of variance of PAD values for the three calibration methods for the three main agricultural activity types

Source of variation	d.f.	s.s.	m.s.	v.r.	F_{pr}
year stratum	4	880.4	220.1	0.65	
year.*Units* stratum					
method	2	1300.8	650.4	1.93	0.161
activity	2	5434.4	2717.2	8.07	0.001
method. activity	4	1062.0	265.5	0.79	0.541
Residual		32	10776.7	336.8	
Total	44	19454.3			

While these results suggested that the GME did not offer an improvement in model's performance, as measured using mean PAD values, compared with the other methods. However, considering only the mean PAD value is misleading. When the variability (i.e. the variance) in model calibration over the 5 year data set is considered, it can be seen that GME offers a much improved inter-annual stability in model's predictions compared with the arbitrary behaviour resulting by using the other two methods. Figure 4.8 shows the PAD values for each calibration method by year within each activity type (Figures 4.8a-c), and the empirical variances of these data (Figure 4.8d). It is apparent that the calibrations achieved by the GME are more stable than those produced by methods which calibrate to a single base year (standard PMP or ME).

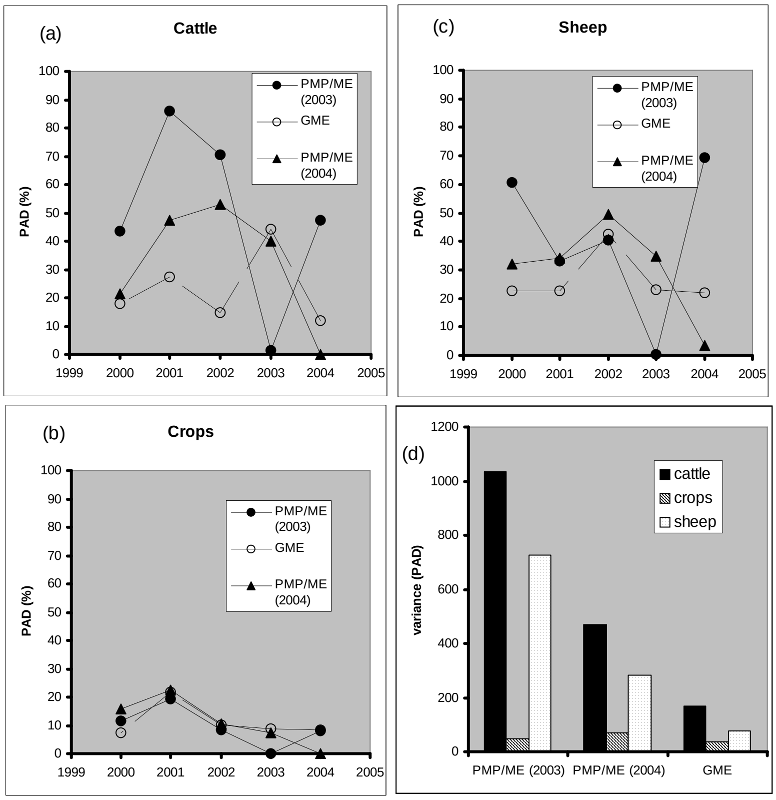

Figure 4.8. PAD values for five years of activity data for three main agricultural activity types for three calibration methods (a-c) and the variance in the PAD values over years (d).

The observed performance of the different calibration methods for the PMP raises a question for modellers about the most appropriate calibration method for their purposes. For applications in which it is important that the PMP model calibrates to (i.e. reproduces) a base year exactly, the standard PMP or ME calibration method might be a better option. However, where the model is expected to consistently reflect the general behaviour of a system, perhaps including the representation of time-averaged behaviour, it is clear that the GME method will produce calibrations in which the risk of very large deviations from observed data is much lower than for methods which calibrate to only to a single year.

4.6. Conclusion

In this chapter the construction of the SASEM model was presented. Specifically, it was shown how, when only a limited amount of data are available, a model which captures the aggregate production decisions of farmers in Scotland as a whole can be developed by using the PMP rationale and the GME econometric estimation method. The model exhibits a number of characteristics that are generally desirable when used as a tool for policy analysis in agriculture. First, some of its objective function coefficients are econometrically estimated. Thus, it is not purely a normative model whose solutions are theoretical prescriptions on how the system should work but rather a descriptive model capable of simulating the system as it really works. Secondly, it does not suffer from drawbacks which are typical to standard LP models such as overspecialisation and step-wise response to changes in the policy conditions (which take the form of changes in prices, technology, or constraints in the system). Furthermore, it is more flexible in responding to these and similar changes because it is not restricted by mechanistically imposed constraints as is quite often the case in LP models. Finally, as demonstrated in section 4.2 the model exhibits lower aggregation bias compared to an LP equivalent (i.e. same spatial and aggregation scale).

Not surprisingly, the SASEM model is not free from strong assumptions and limitations. The model simulates production decisions in a national level and hence does not allow for heterogeneity in behaviour across agricultural producers. Its large spatial resolution also makes the model inappropriate for working in conjunction with biophysical process models, which are often specified at a lower spatial level. In addition, the model covers only a limited number of possible input and output production activities and has a rather simple structure in terms of their interlinkages. Nonetheless, SASEM provides a very useful modelling tool for facilitating the *Follower's* reactions to the *Leader's* decisions within the bi-level policy optimisation problem which is the principal focus of this thesis.

PART II

Multi-objective Bi-level Optimisation of Agricultural Policy: The Leader's model

5. Multi-objective Optimisation

5.1. Introduction

The second part of the Thesis is dedicated to the upper level of the bi-level programming problem of agricultural optimisation. As it was mentioned in the introduction the upper level (or the leader) problem is further complicated by the existence of multiple objective functions. Traditional optimisation methods characterized by a scalar scheme (a single objective function), based on unique solutions search are rather limited in capturing the richness and complexity which solving policy problems frequently presents, because of their requirement of simultaneously balancing multiple objectives. Clearly, an alternative modelling approach that takes into account several different criteria is needed.

This chapter (the first of the two chapters that compose the second part of the thesis) introduces multi-objective decision making (MODM) and the mainstream modelling methods for handling multiple objectives. It is divided into three sections. The first one provides certain fundamental concepts used in multi-objective optimisation programming to help appreciate the methods involved in modelling the agricultural decision processes via the MODM paradigm. Section two presents some of the commonly used classical methods for handling multi-objective optimisation problems. They are referred to as classical in order to be distinguished from the so-called evolutionary methods. The last section of the chapter is devoted to the description of one major group of multi-objective evolutionary optimisation methods, called Multi-objective Genetic Algorithms. The focus of attendance there is on the presentation of the *Elitist Non-Dominated Sorting Genetic Algorithm* (NSGA-II) which is the one that APOLO uses.

5.2. Basics of Multiple Objective Optimisation

According to Zeleny (2005) if the maximum or minimum value of the objective function is set at a predetermined value *a priori*, then it cannot be further optimised (nor maximised or minimised). Similarly, if the value of a constraint is set *a priori*, then it cannot be optimised. Constraints have to become objectives in order to be optimised. Even if the value of the objective is not determined *a priori*, but the constraints are fixed, it still cannot be optimised; it is strictly implied (given) by the constraints.

To discover optimal solutions in multiple criteria decision problems it is necessary to acknowledge multiple concepts of optimality. Multi-objective optimisation applies to an economic problem only when scarce means (constraints) are used to achieve alternative goals. If the means are scarce, but there is only a single goal, then the problem of how to optimally use these means is a technical problem in which case, no value judgement enters into its solution, no balancing is needed and no multi-objective optimisation can take place. A multiple-objective optimisation problem (MOOP) therefore, is about balancing more than one criterion. It has a number of objective functions which are to be minimised or maximised. As in the single-objective optimisation problem, here too the problem usually has a number of constraints which any feasible solution (including the optimal solution) must satisfy. In the following, the multiple-objective optimisation problem in its general form is stated:

$$\begin{array}{lllr} \text{Min/Max} & f_m(\mathbf{x}), & m = 1, 2, \ldots, M & \\ \text{subject to} & g_j(\mathbf{x}) \geq 0, & j = 1, 2, \ldots, J & (1) \\ & h_k(\mathbf{x}) = 0 & k = 1, 2, \ldots, K & \\ & x_i^{(L)} \leq x_i \leq x_i^{(U)}, & i = 1, 2, \ldots, n & \end{array}$$

A solution $\mathbf{x} = (x_1, x_2, \ldots, x_n)^T$ is a vector of n decision variables. An objective function $\mathbf{f(x)} = (f_1(\mathbf{x}), f_2(\mathbf{x}), \ldots, f_m(\mathbf{x}))^T$ is a vector of M objective functions and constitutes an M-dimensional space, called the objective space, in addition to the usual decision variable space. Multi-objective optimisation is sometimes referred to as *vector optimisation* because a vector of objectives, instead of a single objective, is optimised.

For each of the M objectives there exists one different optimum solution, $\mathbf{x}^{*(m)}$ corresponding to an optimum objective value, $f_m^*(\mathbf{x})$. The objective vector $\mathbf{f}^*$ constructed with these individual optimal objective values constitutes the *ideal objective vector*. However, as Shaffer (1985) observed such a Utopian solution where all objectives are simultaneously optimal is unlikely to exist.

Additionally, in problems with more than one conflicting objective, there is no single optimum solution. An optimal set of solutions can exist, but no solution from the optimal set satisfies all objectives better than any other. Since a number of solutions are optimal, in a multi-objective optimisation problem all such solutions are important. The final choice of accepting or implementing one solution involves many other considerations such as the existence of legitimate non-equivalent views of different social actors, or what Munda (2004) calls *social incommensurability*. For example, what constitutes an improvement for one person or society may not be an improvement for another. Individuals or different interest groups having different incentives will prefer more of one attribute if the relative benefits (costs) of that attribute are higher (lower) (Yiridoe and Weersink, 1997).

Decision makers can only make rational choices or tradeoffs among alternatives based on available information on the relative values of available choices. The value of a tradeoff between two attributes shows how much of one attribute has to be foregone (sacrificed) to attain a per unit improvement in the other. In other words, the tradeoff curve can be thought of as a transformation curve where an attribute can be measured in terms of foregone units of another attribute. In economics this is the principle of opportunity cost which converts non-economic attributes into monetary values. The information provided by the tradeoff curve would support a bargaining process among different interest groups by providing various alternatives in order for a compromise solution to be agreed upon. Figure 5.1 shows the role of tradeoff curves in the process of MODM. Tradeoffs can be obtained by using the concept of dominance as explained in the next section.

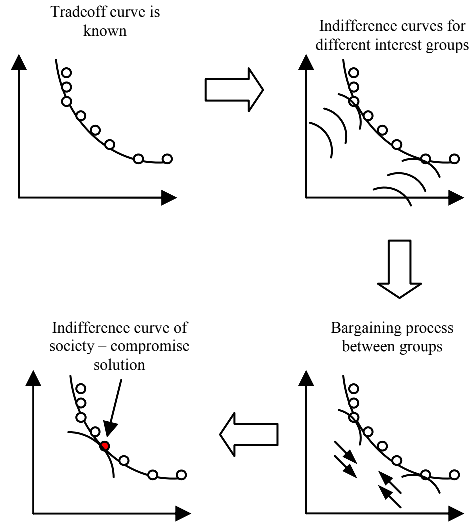

Figure 5.1. Tradeoffs within the process of multiobjective decision making

5.2.1. Dominance and Pareto Optimality

The concepts of dominance and Pareto optimality play a vital role in traditional economic theory and are also fundamental ideas within the MODM paradigm, as all the optimisation approaches within this paradigm use these concepts to search for *non-dominated* or *Pareto-optimal* solutions. In their algorithms, two solutions are compared on the basis of whether one dominates the other solution or not with respect to all the objectives under consideration. In order to cover both minimisation and maximisation of objective functions, the operators $\prec$ and $\succ$ between two solutions *i* and *j* are used. For instance, $i \prec j$ denotes that solution *i* is better than solution *j* on a particular objective. Similarly, $i \succ j$ for a particular objective implies that solution *i* is worse than solution *j* on this objective. Now, a solution x^1 is said to dominate another solution x^2, if both of the following two conditions are true:

1. The solution x^1 is no worse than x^2 in all *M* objectives, or $f_j(x^1) \succ f_j(x^2)$ for all j = 1,2,…, *M*.
2. The solution x^1 is strictly better than x^2 in at least one objective, or $f_{\hat{j}}(x^1) \prec f_{\hat{j}}(x^2)$ for at least one $\hat{j} \in \{1, 2,…, M\}$.

If any of the above conditions is violated, the solution x^1 does not dominate the solution x^2. If x^1 dominates the solution x^2, the following relationship expressions hold: x^2 is dominated by x^1; x^1 is non-dominated by x^2; x^1 is non-inferior to x^2.

Figure 5.2 plots a number of different solutions to a two-objective optimisation problem where the objective function 1 needs to be maximised while the objective function 2 needs to be minimised. The above definition of domination can be used to decide which solution is better among any two given solutions with respect to both objectives. For instance, solution 1 is better than solution 2 in both objectives. Thus, both of the conditions of domination are satisfied and therefore solution 1 dominates solution 2.

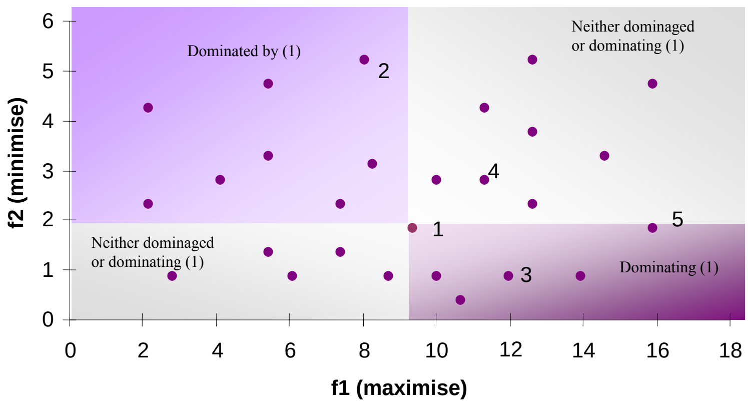

Figure 5.2. Dominance & non-dominance of solutions

However, when comparing solution 1 to solution 5 it can be seen that solution 5 dominates solution 1 because it is better in the first objective and no worse (in fact, they are equal) in the second objective. For a given set of feasible solutions (search space) after performing all possible pair-wise comparisons a set of solutions is expected to be found where any two solutions do not dominate each other. This set is called the *non-dominated* or else *Pareto-optimal* set and the solutions that belong to this set are called non-dominated, Pareto-optimal or efficient solutions. These are feasible solutions such that no other feasible solution can achieve the same or better performance for all the criteria under consideration and strictly better for at least one criterion. In other words, a Pareto-optimal solution is a feasible solution for which an increase in the value of one criterion can only be achieved by degrading the value of at least one other criterion.

A large number of multi-objective optimisation programming (MOOP) methods have been developed in order to find the Pareto-optimal set. They are divided into two broad categories: *Classical* and *Evolutionary*. The distinction is based on the type of search procedure used and on whether one or a set of optimal solutions are found. Classical methods have algorithms that use single solution update in every iteration and deterministic search and optimisation procedures. Evolutionary methods process a population of solutions in each iteration and their search and optimisation procedures mimic natural evolutionary principles.

5.3. Classical Methods of Multi-objective Optimisation

Classical methods, collectively known as multi-criteria decision making, have been around for at least the past four decades. During this period, many algorithms have also been suggested. Many researchers have attempted to classify algorithms according to various considerations. Hwang *et al.,* (1979) and later Miettinen (1999) classified them into the following four classes: (1) No-preference methods, (2) *Posteriori* methods, (3) *A priori* methods and, (4) Interactive methods.

On the one hand, algorithms that belong in the first class do not assume any preference information about the importance of objectives; they do not search for multiple Pareto-optimal solutions, rather, a heuristic is used to find a single optimal solution. In contrast, *posteriori* methods use some information about preference between objectives and require some knowledge on algorithmic parameters to iteratively generate a set of Pareto-optimal solutions. *A priori* methods use as much information about the preferences of objectives as possible and usually find a single preferred Pareto-optimal solution. Finally, interactive algorithms use the preference information progressively through an interaction between the decision maker (DM) and the results generated at various stages during the run of the algorithm.

An outline of classical methods in the order of increasing use of preference information is presented in the following paragraphs.

5.3.1. The weighting method

The basic idea here, as the name suggests, is to "scalarise" the set of all objectives into a single objective function by pre-multiplying each objective function with a weight before they are added. Adding up the objective functions also depends on the scaling of each objective because usually different objectives take different orders of magnitude. When such objectives are combined to form a composite objective function, it would be better to *normalise* them by scaling them appropriately so that each has more or less the same order of magnitude. After the objectives are normalised, a composite objective function $F(\mathbf{x})$ can be formed by summing the weighted normalised objectives and a MOOP problem with Q objectives to be minimised or maximised can be converted to a single-objective optimisation problem with the following mathematical programming structure:

$$\begin{array}{lll} \text{min or max} & \mathrm{F}(\mathbf{x}) = \sum_{q=1}^{Q} w_q f_q(\mathbf{x}), & \\ \text{subject to} & g_j(\mathbf{x}) \geq 0 & j = 1, 2, \ldots, J \\ & \mathrm{h}_\mathrm{k}(\mathbf{x}) = 0 & \mathrm{k} = 1, 2, \ldots, \mathrm{K} \end{array}$$

Since the optimal solution of the above problem does not change if all weights are multiplied by a constant, it is the usual practice to choose positive weights such that their sum is one, or $\sum_{q=1}^{Q} w_q = 1$. Subsequently, the Pareto-optimal set is generated through parametric variation of weights (Zadeh, 1963). The weight of an objective usually reflects the objective's relative importance (preference) in the problem at hand (Deb, 2001). However, the interpretation of the weights is correct only if the objective function stated for the MOOP problem truly corresponds to the utility function for the decision maker. In a general way, the weighting sum method treats the weights only as parameters that can be varied systematically to generate the efficient (Pareto-optimal) set (Romero and Rehman, 2003). Figure 5.3 illustrates how

the weighting approach can find the efficient set. For simplicity a problem of minimising two objectives is considered.

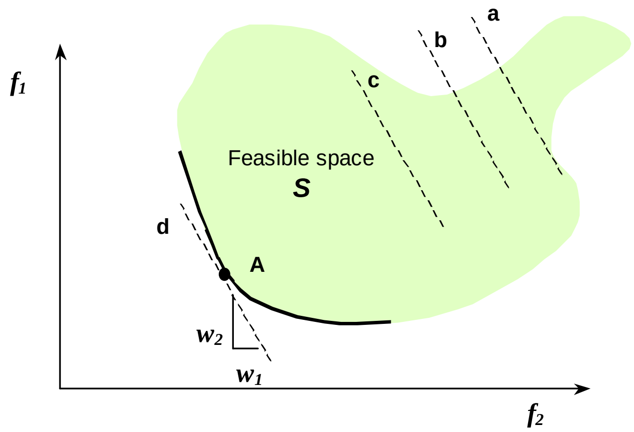

Figure 5.3. Illustration of the weighting method on a convex efficient set

Since the objective function F is a linear combination of both objectives f_1 and f_2, the contour surface of F on the objective space is a straight line whose slope is related to the choice of the weight vector. The location of the line depends on the value of F i.e. the effect of lowering the contour line form ‘a’ to ‘b’ is in effect jumping from solutions of higher F values to a lower one. The minimum solution corresponding to the weight vector therefore is the point ‘A’ where the contour line ‘d’ is tangential to the objective space. By using different weight vectors the slope of the contour line changes resulting in different optimal solutions. The Pareto-optimal set is found after the problem is solved a number of times for multiple weight vectors.

5.3.2. The constraint method

This method reformulates the MOOP by just optimising one of the objectives and restricting the rest within user-specified values in order to generate the efficient set. The modified *m*-objective problem is as follows:

$$\begin{array}{lll} \text{Min or max} & f_\mu(\mathbf{x}) & \\ \text{Subject to} & f_q(\mathbf{x}) \le \in_q & q = 1, 2, \ldots, Q \text{ and } q \ne \mu \\ & g_j(\mathbf{x}) \ge 0 & \\ & h_k(\mathbf{x}) = 0 & \end{array}$$

where the parameter $\in_1^a$ represents an upper bound of the value of f_μ. Through parametric variations the efficient set is generated (Figure 5.4). The constraint method guarantees efficient solutions only when the parametric constraints are binding (points 'B' and 'C') in the optimal solutions.

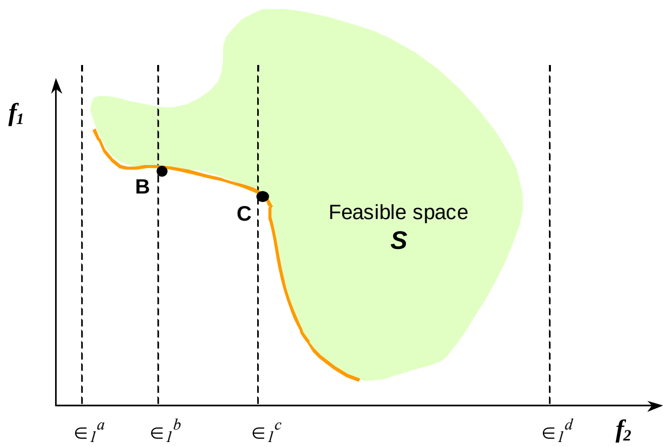

Figure 5.4. The constraint method

5.3.3. Goal Programming methods

Goal programming which is perhaps the oldest multiple criteria decision making technique gained popularity especially after the work of Lee (1972) and Ignizio (1978). Romero and Rehman (2003) provide the following definition of goal programming:

> '…its main aim is a simultaneous optimisation of several goals, by minimising the deviations from the desired targets for each of the objectives and what is actually achievable in relation to the targets set.'

There can be four different types of goals: Less-than-equal-to, greater-than-equal-to, equal-to and, within a range. In order to achieve the goals, two non-negative deviation variables (n and p) are usually introduced that convert inequalities into equalities. For the less-than-equal-to type goal, the positive deviation p is subtracted from the objective function, so that $f(x) - p \le t$. For the greater-than-equal-to type goal, a negative deviation n is added to the objective function, so that $f(x) + n \ge t$. For the equal-to type goal, the objective function needs to have the target value t, and thus both positive and negative deviations are used, so that $f(x) - p + n = t$. The forth type of goal is handled by using two constraints: $f(x) - p \le t^{lower}$ and $f(x) + n \ge t^{upper}$. All of the above constraints can be replaced by a generic equality constraint:

$$f(x) - p + n = t$$

Goal programming methods, each being a specific variant of GP, differ in the ways that the deviations are minimised. The best known and widely used variants are briefly discussed below

5.3.3.1. Weighted Goal Programming

In this variant of GP a composite objective function which is the sum of all the weighted deviations among the goals and the aspiration levels is minimised. The following problem is constructed:

$$\begin{array}{lll} \text{Min} & \sum_{g=1}^{G}(\alpha_g p_g + \beta_g n_g) & \\ \text{Subject to} & f_g(x) - p_g + n_g = t_g & g = 1, 2,\ldots, G \\ & X \in S & \\ & n_g, p_g \ge 0 & g = 1, 2,\ldots, G \end{array}$$

where the parameters α_g and β_g are weighting factors for positive and negative deviations of the g-th objective. Usually, the weights are fixed according to the relative importance attached to each goal by the decision maker, which makes the method subjective to the user. This matter is illustrated through the following simple example found in Deb (2001) and reproduced here:

goal $\quad f_1 = 5x_1 \leq 10$

goal $\quad f_2 = \dfrac{10+(x_2-6)^2}{10x_1} \leq 0.2$

subject to $\quad S \equiv (1 \leq x_1 \leq 10,\ 0 \leq x_2 \leq 10)$

with a GP problem as follows:

$$\begin{aligned} \min \quad & \alpha_1 p_1 + \alpha_2 p_2 \\ \text{s.t.} \quad & 5x_1 - p_1 \leq 10 \\ & \frac{10+(x_2-6)^2}{10x_1} - p_2 \leq 0.2 \\ & 1 \leq x_1 \leq 10, \quad 0 \leq x_2 \leq 10, \quad p_1, p_2 \geq 0 \end{aligned}$$

From Figure 5.5 it is clear that there exists no feasible solution that achieves both goals. The GM method finds the solution which violates either or both of the goals in a minimum sense, namely the solution that minimises the deviation from the target space (shown on the right) in both objectives. There exist many such solutions each one corresponding to different set of weights α_1 and α_2. Solutions A, B and C shown on figure 5.3 are the optimal solutions found by the GP for $\alpha_1 = \alpha_2 = 0.5$, $\alpha_1 = 1$ and $\alpha_2 = 0$, and $\alpha_1 = 0$ and $\alpha_2 = 1$ correspondingly.

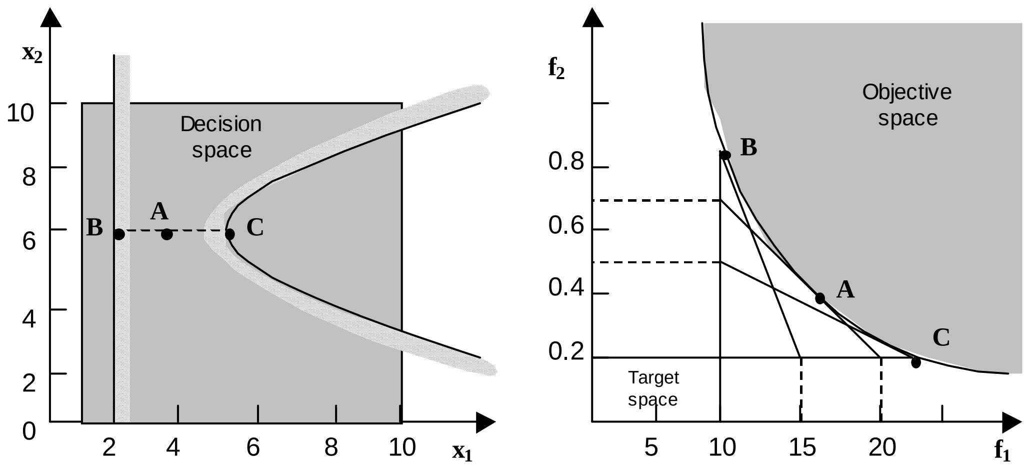

Figure 5.5. The GP problem shown in solution space (left) and in objective space (right) with contours produced by the objective function α_1 p_1 + α_2 p_2.

5.3.3.2. Lexicographic Goal Programming

In this approach different goals are categorised into several level of pre-emptive priorities. In other words, goals of higher priority are immeasurably preferable of goals situated in lower priorities. Hence, LGP formulates and solves a number of sequential goal programming problems. First the model is solved by minimising the vector of deviations from the targets of the goals having priority one. This vector is called the achievement function. If multiple solutions are found, another GP model is formulated with goals having the second-level priority and solved by minimising any deviations in the goals of priority two. The goals of first-level priority are used as hard constraints so that the obtained solutions does not violate the goals of priority one. This process continues with goals of the next higher-level priorities in sequence and terminates as soon as a single solution is found.

5.3.4. Limitations of the classical methods

In terms of finding multiple Pareto-optimal solutions, classical multi-objective optimisation algorithms may encounter a number of difficulties. First, only one Pareto-optimal solution is found in any single run of these algorithms. Such a solution is specific to the parameters used to convert the multi-objective optimisation problem into a single-objective optimisation one. In order to find a different Pareto-optimal solution, the parameters must be changed and the resulting problem has to be solved again. Second, in multi-objective optimisation problems with a nonconvex Pareto-optimal set, not all members of that set can be found by some algorithms, and third, all require some knowledge about the problem such as suitable weights or goal values. This is because a uniformly spaced set of weight vectors may not produce uniformly spaced Pareto-optimal solutions, and different weights may produce identical solutions.

Despite these shortcomings, classical methods have a number of advantages. Apart from being simple and easy to implement on a computer, the proofs of convergence to the true Pareto-optimal set makes them very appealing in solving real-world problems.

5.4. Multi-objective Evolutionary Optimisation Methods

Another branch of methods for handling multi-objective optimisation problems are the multi-objective evolutionary algorithms. An evolutionary algorithm (EA) indicates a population-based metaheuristic optimisation algorithm that uses mechanisms inspired by biological evolution, such as reproduction, mutation, recombination, natural selection and survival of the fittest. Candidate solutions to the optimisation problem play the role of individuals in a population, and the cost function determines the environment within which the solutions "live". Evolution of the population then takes place after the repeated application of the above operators. Their main steps are: initialisation, evaluation, selection and variation. These steps are repeated until a stopping criterion is reached (see Algorithm 1). EA are robust search algorithms and require minimal user-provided information. One of their main advantages is that they are able to approximate the Pareto-optimal set in a single optimisation run.

Algorithm 1: A pseudocode for evolutionary algorithms

```
t:= 0;
initialise population P(t);
initial evaluation P(t);
while stopping criteria not met do
    t:=t+1;
    select P(t) from P(t-1);
    change P(t);
    evaluate P(t);
end
```

EA can be divided into three broad categories: Genetic Algorithms (GA), Evolution Strategies and Evolutionary Programming. Most of these categories have similarities, but differ in the details of their implementation and the nature of the particular problem to which they have been applied. The three variants differ in (1) the data structure that they apply, and (2) the values on which several parameters of the algorithm are set on during the optimisation process. Genetic algorithms use a fixed-length vector of a discrete alphabet with low cardinality, Evolution strategies use a variable-length vector of real values, and Evolutionary programmes employ tree structures of variable length.

As in classical optimisation, early researchers of evolutionary algorithms have also realised the need of extending single-objective EA to handle multiple objectives. Most multi-objective evolutionary algorithms modify single-objective evolutionary algorithms in special ways. The first real implementation of a multi-objective evolutionary algorithm was suggested by Schaffer (1984) who applied the *dominance* relation between solutions to find a set of non-dominated solutions. He called his GA the *vector evaluated GA*, or VEGA. VEGA evaluated an objective vector (instead of a scalar objective function) with each element of the vector representing an objective function. Goldberg (1989) realised a better implementation of domination principle in an EA and suggested a new non-dominated sorting procedure for assigning more copies to non-dominated solutions in a population. Based on Goldberg's suggestions, several independent groups of researchers have developed different versions of multi-objective EA.

5.4.1. Multi-objective Genetic Algorithms

Over the last decade Multi-objective genetic algorithms (MOGA) have been extensively used as search and optimisation tools in various multi-objective optimisation problems. This section serves as a background for the description of the algorithm used to solve the multi-objective bi-level optimisation problem treated in the present study. It has been developed by Deb *et al*, (2000) and is called *Elitist Non-Dominated Sorting Genetic Algorithm* or NSGA-II. As a starting point a description of the general working principles of binary-coded multi-objective genetic algorithms is given.

5.4.1.1. Binary Representation of the Decision Variables

The important starting point for a GA is a pseudochromosomal representation of a candidate solution that uses a constant length chromosome string. The variable represented by a position in the string is analogous to a *gene*, and its value is known as an *allele*. The most popular encoding is a binary coding, where the alleles take a value of 0 or 1 (see Figure 5.6). This involves a mapping mechanism between the solution space (real numbers) and the chromosome string (binary representation of fixed length) shown in Figure 5.7. The precision of the representation can be improved by increasing the number of bits in the binary string.

$$\underbrace{110001010011101}_{X_1}\ \underbrace{011110010001001}_{X_2}$$

Figure 5.6. A binary string of two variables with 15-bit encoding

5.4.1.2. Population Initialisation

The initialisation of the population specifies the starting points of the search. The goal is to create an initial population with a good coverage of the search space, and thereby have a diverse gene pool with potential to explore as much of the search space as possible. The initial population can be created in a number of ways. One option is to randomly assign chromosomes using a uniform distribution. In an alternative setup, chromosomes can be deterministically scattered evenly over the whole search space according to a regular grid-layout. A third alternative is to incorporate expert knowledge about the objective function. After choosing a string representation scheme and creating a population of strings, the genetic operations to such strings can be applied to find better solutions.

To carry out the transformation a real number, r, is first transformed to a base-10 integer, z using:

$$r = mz + c \qquad (1)$$

This integer is then converted to the binary string. The values of m and c depend on the location and width of the search space. Expressions for m and c can be derived from the two simultaneous equations;

$$r_{min} = mz_{min} + c \qquad (2a)$$

$$r_{max} = mz_{max} + c \qquad (2b)$$

where r_{min}, r_{max}, z_{min} and z_{max} represent the minimum and maximum possible parameters in real and integer representations respectively. The smallest binary number that can be represented is of the form *000…0* which equates to 0 in base-10, so $z_{min} = 0$. z_{max} is given by:

$$z_{max} = 2^l - 1 \qquad (3)$$

where l is the length of the binary string used. Substracting (2a) from (2b) gives:

$$r_{max} - r_{min} = m(z_{max} - z_{min})$$

or

$$m = (r_{max} - r_{min}) / (z_{max} - z_{min})$$

Applying (3) and remembering that $z_{min} = 0$ gives:

$$m = (r_{max} - r_{min}) / (2^l - 1) \qquad (4)$$

Finally, rearranging (2a) yields:

$$c = r_{min} - mz_{min}$$

or (as $z_{min} = 0$)

$$c = r_{min} \qquad (5)$$

Equations (1), (4) and (5) then define the required transformation:

$z = (r - r_{min})\ (2^l - 1) / \ (r_{max} - r_{min})$

Figure 5.7. Binary encoding of real numbers

5.4.1.3. Assigning Fitness to a Solution

Each solution is tested empirically in an "environment" and is assigned a numerical evaluation of its merit by a fitness function. The environment can be almost anything and in a MOOP problem it represents the attributes reflecting the DM's values related to an objective reality. These attributes can be measured and expressed as mathematical functions of the decision variables and comprise the objective functions of the problem. MOGAs consider the objectives as directions of improvement of one or more of the attributes of a successful or "fit" solution, implying a sense of either "the more of the attribute, the fitter the solution" or "the less of the attribute, the fitter the solution". However, the MOGA needs only one fitness value for evaluating and sorting the solutions. Consequently, a technique has to be used to find a single metric from a number of objective functions. As already mentioned, Schaffer (1984) and later Goldberg (1989) suggested the implementation of the domination concept for that purpose. The main principles of a domination-based sorting are presented below.

5.4.1.4. Non-Dominated Sorting of a Population of Solutions

One of the most critical stages of a MOGA is the assignment of fitness values to solutions in the face of multiple objectives. In order to evaluate the solutions of a population with respect to multiple objectives, MOGA use the dominance relations to identify the better (non-dominated) of two given solutions. Many algorithms for finding the non-dominated set from a given population of solutions have been suggested. Here, first the naive and slow approach and then the non-dominated sorting as described by Deb (2001, pp. 34-43) are presented.

Naïve and Slow

Each solution *i* is compared with every other solution in the population to check if it is dominated by any solution. If solution *i* is found to be dominated by any solution, it cannot belong to the non-dominated set. If, on the contrary, no solution is found to dominate solution *i* then the latter is a member of the non-dominated set. The following pseudo algorithm (Algorithm 2) describes the procedure of finding the non-dominated set in a given population P of size N.

Algorithm 2: A pseudocode for naïve and slow sorting	
Step 1	Set population counter $i = 1$ and create an empty non-dominated set P'.
Step 2	Set another population counter $j = 1$. For j ($j \in P$) = 1 To N if $j \neq i$ check if j dominates solution i. If yes, go to Step 4.
Step 3	Set $P' = P' \cup \{i\}$.
Step 4	Increment i by one. If $i \leq N$, go to Step 2; otherwise stop and declare P' as the non-dominated set.

Non-Dominated Sorting

The naïve and slow procedure finds only the best non-dominated front in a population. Nonetheless, there exist some MOGA which require the entire population to be classified (sorted) into all the possible non-domination levels (classes or fronts) according to an ascending level of non-domination. Solutions belonging to front 1 are regarded as the best non-dominated solutions. These are disregarded from the initial population and the new non-dominated solutions of the remaining population are then found and are nominated non-dominated solutions of level 2. This procedure is continued until all members of the set are sorted into a non-dominated class. It is important to reiterate that non-dominated solutions of level n are better than non-dominated solutions of level n+1. The steps of the non-dominated sorting of a population are described below (Algorithm 3).

Algorithm 3: A pseudocode for non-donimated sorting
Step 1 Set non-domination level counter j = 1.
Step 2 Create an empty non-dominated set P_j.
Step 3 Use an algorithm (such as the one described earlier) to find the non-dominated set P′ of population P
Step 4 Update $P_j = P'$ and $P = P \setminus P'$.
Step 5 If $P \neq \emptyset$, increment j by one and go to Step 2. Otherwise, stop and declare all non-dominated sets P_j.

Figure 5.8 shows an example of non-dominated sorting of a population of 11 solutions into different non-dominated fronts (indicated by the dashed lines) in a two-objective maximisation-minimisation problem. It is clear that all solutions in the first front are the best solutions. The second best solutions in the population are those that belong to the second front and so on. However, any two members from the same front cannot be said to be better than one another with respect to both objectives. The total number of fronts (or classes) depends on the population and the underlying problem.

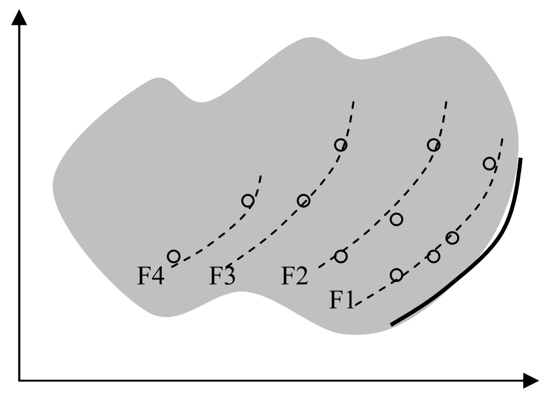

Figure 5.8. Solutions classified into various non-dominated fronts

By examining Figure 5.8 carefully it is apparent that solutions of the first front are best in terms of their closeness to the true Pareto-optimal set indicated by the solid line. Thus, it is expected the highest fitness to be assigned to solutions of the best non-dominated front. In this way, solutions in a population are ranked in relation to their fitness values and the *selection* operator can use this information to select the solution that will enter the mating pool.

After all solutions of the population have been evaluated, a termination condition is checked. If it is not satisfied, the solutions are modified by three main genetic operators and a new (and hopefully better) population is created. The *generation counter* is incremented to indicate that one generation (or one iteration in the parlance of classical search methods) of the GA is completed. Through repeated selection the individual fitness as well as the total fitness of the population as a whole is increased as a result of a probabilistic process that favours the reproduction of highly fit solutions as new generations of solutions are produced. After a number of generations the fittest solutions are found.

5.4.1.5. Selection Operator

Once all solutions have been assigned a fitness value the selection operator uses those values as the basis for identifying good solutions and selecting (by making duplicates of good solutions and eliminating bad solutions) the solution that will be placed in the mating pool. The mating pool refers to the population of only those solutions that are chosen for mating. Individuals (solutions) that do not enter the mating pool are removed. Selection is equivalent to the procedure of exploitation of good solutions. It essentially defines how the algorithm updates the population from one iteration to the next. There exist a number of selection methods. *Tournament selection*, *proportionate selection* and *ranking selection* (Goldberg *et al.*, 1991) are commonly used.

In the *tournament selection*, each solution is made to participate in exactly two pair-wise tournaments and the winner solution is chosen and placed in the mating pool. The best solutions in a population will win both tournaments thereby two copies of it are made in the mating pool. In the same way, the worst solutions will lose both times and will be discarded from the mating population. At the end of all tournaments any solution in a population will have zero, one or two copies in the mating pool. Tournament selection is probably the most commonly used because it is easy to implement, produces good results within short time and is controlled by only a small number of parameters.

In the *proportionate selection* method copies of solutions are made. The number of copies is proportional to the solution's fitness value. This selection operator can be thought of as a roulette-wheel mechanism, where the wheel is divided into a number of divisions equal to the population size, N, where the size of each division is marked in proportion to the fitness of each population member. Thereafter, the wheel is spun

N times and each time the solution indicated by the wheel pointer is placed in the mating pool. The problem with this selection method is that the outcome is dependent on the true value of the fitness instead of the relative fitness values of the population members. This "scaling" difficulty can be avoided if the solutions are first sorted according to their fitness, from the worst (rank 1) to the best (rank N). Then, each member in the sorted list is assigned a fitness value equal to its rank. Thereafter, the proportionate selection operator is applied with the ranked fitness values. This method is known as the *ranking selection* operator.

5.4.1.6. Crossover Operator

The term "crossover" is used to refer to the exchange of homologous substrings between individuals, although the biological term "crossing over" implies exchange between chromosomes within an individual organism. The primary aim of the crossover operator is to create the "offspring" solution from the members (called parent solutions) of the mating pool. The creation of new solutions starts by picking two strings from the mating pool at random. Crossover exchanges some portion(s) of the strings to create two new offspring solutions. In a *single-point crossover* this is performed by randomly choosing a crossing site along the string, and by replacing all bits on one side of the crossing site with the corresponding part on the other string (see Figure 5.9).

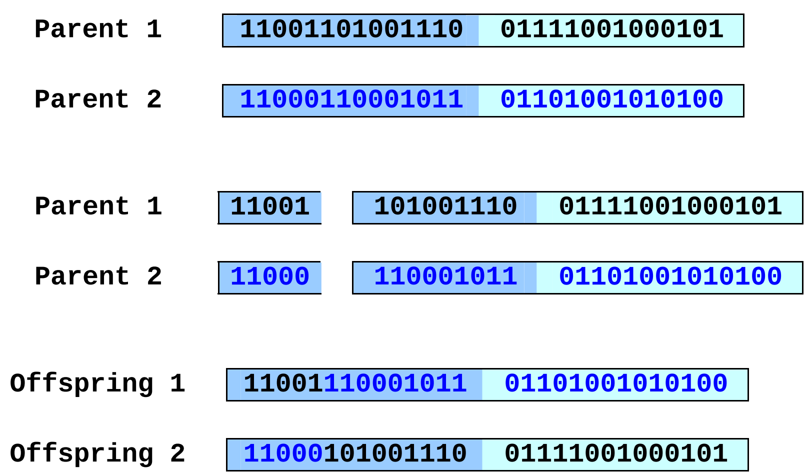

Figure 5.9. One-point crossover operator

The same concept of exchanging partial information between two strings can also be achieved with more than one crossover site. In this case the crossover operator is called *N-point crossover*. In order to preserve some good solutions selected during the selection operator, a stochastic sorting routine is used that determines the proportion of the strings used in the crossover operation and those simply copied to the new population. Creating new strings every next generation is equivalent to the procedure of exploring the feasible space for optimal solutions.

5.4.1.7. Mutation Operator

If only the crossover operator is used to produce offspring, one potential problem that may arise is that if all the chromosomes in the initial population have the same value at a particular position then all future offspring will have this same value at this position. For example, if all the chromosomes have a 0 in position two then all future offspring will have a 0 at position two. To combat this potentially undesirable situation and to further enhance the diversity of solutions in every generation a bit-wise mutation operator is used. Mutation attempts to introduce some random alteration of the genes by flipping a 0 to 1 and vice versa with a mutation probability of p_m. Typically this occurs infrequently so mutation is of the order of about one bit changed in a thousand tested. Each bit in each chromosome is checked for possible mutation by generating a random number between zero and one. If this number is less than or equal to the given mutation probability p_m then the bit value is changed.

5.4.1.8. Elite-Preserving Operator

The elite-preserving operator preserves and uses previously found best solutions in subsequent generations. In a simple implementation the best ε% of the current population is directly copied to the next generation. The rest (100 - ε%) of the new population is created by the usual genetic operations applied on the entire current population (including the chosen ε% elite members). This guarantees not only that the best solutions are included in the new generation, but that they also participate in the mating process for creating new population members.

5.4.2. Elitist Non-Dominated Sorting Genetic Algorithm (NSGA-II)

This section presents the *Elitist Non-Dominated Sorting Genetic Algorithm* or NSGA-II for short, developed by Deb *et al* (2000), which is the MOGA used in the APOLO model. It draws heavily on the description given in Deb (2001 pp 245-9). NSGA-II uses an elite-preservation strategy and an explicit diversity-preserving mechanism. The algorithm can be outlined in four major steps.

In step one all N solutions of the population $\mathbf{P_t}$ (the subscript t denotes the generation) are used to create equal number of offspring (population $\mathbf{O_t}$) using genetic operators. The two populations are combined together to form a population of size $2N$. At this stage, the non-dominated sorting is used to classify the solutions into different fronts. A temporary population $\mathbf{T_t}$ is created first by solutions of the best non-dominated front and continues with solutions of the second best front, followed by the third and so on. However, only N solutions will be allowed to enter population $\mathbf{T_t}$. Therefore, in step two, fronts accommodating solutions ranked from N+1 to 2N are simply deleted. When the front accommodating the N^{th} solution is considered all solutions from the N^{th} solution and up are discarded. The rejection of the discarded members is done in step three by using a *niching strategy*. This procedure chooses the members of a front, which reside in the least crowded region (niche) in that front. For this purpose NSGA-II uses a *crowding distance metric*, which is described later below. The front with the solutions left is the least good front of the temporary population $\mathbf{T_t}$. This scenario is illustrated in Figure 5.10.

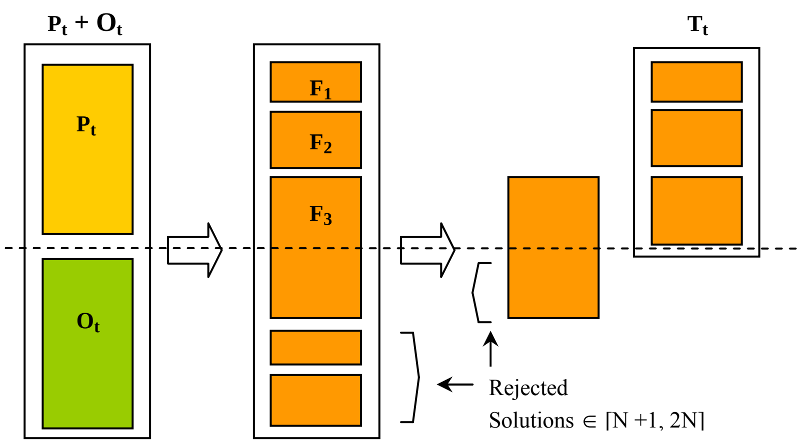

Figure 5.10. Schematic representation of the NSGA-II

In the last step the mating pool is created from the temporary population $\mathbf{T_t}$ by using the selection operator. Then the offspring population $\mathbf{P_{t+1}}$ (next generation), is created by applying the crossover and mutation operators. The selection operator that the NSGA-II uses is called the *crowded tournament selection* and is described in the subsection that follows. For crossover and mutation NSGA-II uses standard operators described previously.

5.4.2.1. Crowded Tournament Selection Operator

This is a selection operator which also ensures that a diverse set of solutions will be chosen for mating. It compares two solutions with respect to two attributes and returns the winner of the tournament. The two attributes are:

1. A non-domination rank in the mating pool
2. A local crowding distance (which is described in the next subsection)

A solution wins the tournament if either it has a better rank or if it has the same rank but has a larger crowding distance than the other solution. All winner solutions enter the mating pool that will undergo single-point crossover and bit-wise mutation to give the offspring population (next generation). It should be made clear that the same solution can be represented in the mating pool for as many times as the number of tournaments it has won. Thus, good solutions have more copies in the mating pool than less good solutions.

5.4.2.2. Crowding Distance

When the number of solutions that belong to the last non-dominated front included in the temporary population $\mathbf{T_t}$ exceeds the number of the remaining spaces in $\mathbf{T_t}$ an operator is applied to discard the superfluous solutions. The mGA should maintain a diverse population in order to prevent premature convergence and to achieve a well distributed trade-off (Pareto-optimal) front. To ensure that solutions do not form clusters by chance, NSGA-II uses a crowding distance assignment procedure to measure the density of solutions surrounding a particular solution *i* to arrange the population in descending order. This measurement, called *crowding distance*, d_I gives an estimate of the average distance of two solutions on either side of a solution along each of the objectives. A crowding distance $d_{I_j^m}$ along an objective *m = (1,…, l)* is given by the following formula:

$$d_{I_j^m} = \frac{f_m^{(I_{j+1}^m)} - f_m^{(I_{j-1}^m)}}{f_m^{\max} - f_m^{\min}} \quad \forall j = 2,\ldots,(l-1)$$

$$d_{I_j^m} = \infty \ when \ j = 1, l$$

I_j denotes the *j*-th solution, and I_1 and I_l denote the lowest and upper most boundary solutions when the set of solutions is sorted in worse order of f_m. The parameters $f_m^{\max}$ and $f_m^{\min}$ can be set as the population maximum and minimum values of the m-th objective function. The sum of distances over one objective is therefore given by:

$$d_{I_j} = \sum_{j=1}^{l} d_{I_j^m}$$

d_{I_j} estimates the perimeter of the cuboid formed by using the nearest neighbours as the vertices. In Figure 5.11, the crowding distance d_I of the *j*-th solution I_j in its front is the average side length of the cuboid (shown by the dashed box) or the sum of the two distances $d_{I_j^1}$ along objectives 1 and 2 respectively. By selecting solutions on the basis of large crowding distances, the algorithm avoids the danger of selecting clusters of very similar solutions and maintains diversity within the population.

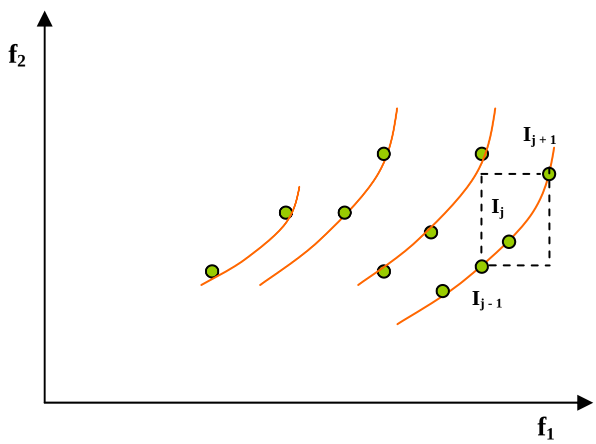

Figure 5.11. The crowding distance calculation

5.4.2.3. Constraint Handling

A common way of handling constraints in evolutionary algorithms is by adding a penalty proportional to the constraint violation to the objective value. The constraint handling method discussed here, which was implemented by Deb (2001), is based on penalty functions where an exterior penalty term which penalises infeasible solutions is used. The method uses the same tournament selection operator as described above, in which two solutions are picked from the population and the better solution is selected. However, in the presence of constraints a change is required in the definition of the domination concept which now becomes *constraint-domination*. Each solution can be either feasible or infeasible. When both solutions are infeasible the one with smaller overall constraint violation *constraint-dominates* the other and is selected (wins). In case where only one is feasible it is intuitive that the feasible solution is the *non-constraint-dominated*. When both solutions are feasible, the one that belongs to the better non-dominated front is also the *non-constraint-dominated* solution and wins the tournament. If they belong to the same non-dominated front the crowding criterion (which maintains solution diversity) can be applied to resolve the tie. Hence, the solution which belongs to the least crowded region in that non-dominated front is chosen.

5.4.3. Salient Issues of Multi-objective Genetic Algorithms

There are several issues related to the design development and application of MOGAs. All are more or less related to the two main tasks that a MOGA should accomplish in solving multi-objective optimisation problems. First, the algorithm should guide the search towards the global Pareto-optimal region and, second, it should maintain population diversity (in the objective space, variable space, or both) in the non-dominated fronts (Deb, 1999). This section is concerned with those that are the most important in the implementation of the MOGA in the multi-objective bi-level optimisation problem studied in the thesis.

5.4.3.1. Number of objectives

In multi-objective optimisation the problem difficulty varies with the number of objectives. When the number of objectives increases, the dimensionality of the objective space and the Pareto-optimal surface also increases. The task of the mGA is to reach this surface from the feasible search space and distribute solutions as uniformly as possible over the surface. It is also intuitive that the number of non-dominated solutions in the initial random population will also increase with an increase in number of objectives. One way to calculate the number, n, of non-dominated solutions in a randomly created population is to calculate the probability $p(k)$ of having a population with exactly $K \in [1, N]$ non-dominated solutions. The obtained probability distribution gives the expected value. However the computation of probabilities, for reasons explained in Deb (2001, pp 416-7) is difficult. Experiments with different number of objectives and population size gave a growth of n similar to a logistic growth pattern. In other words as the number of objective functions increases, most solutions in the population belong to the non-dominated front. With respect to the population size, as it increases the proportion of non-dominated solutions decreases (for a fixed number of objectives). Most MOGAs assign the same (or similar) fitness to all solutions of the same front. When this happens, there is no selection advantage to any of these solutions. Only the crossover and mutation operators can create solutions in a better front whereby the search can proceed towards the Pareto-optimal region. Deb (2001) suggests two solutions to this problem; to use large population size and/or to modify the algorithm.

5.4.3.2. Parameter Setting

Every part of a GA has parameters. There are parameters for the population, the representation, and the evolutionary operators. GAs are sensitive to their parameters. For instance, the probability of mutation, p_m, can have quite an impact on the performance of the algorithm. A fundamental problem in determining the best parameters is that they depend on the problem. One set of parameters may yield good performance for some problems but less good performance on other problems. Hence, no set of parameters is superior on all problems. The sensitivity of GAs to parameter values has lead to a wide variety of parameter control approaches which can be divided into two types depending on whether not the control method adapts to the search process. Non-adaptive control methods keep the parameter values constant or changed by a simple parameterised function of the generation counter. Fortunately, this will not decrease the performance of the algorithm significantly. The main drawback is the significant amount of manual tuning involved, and that the values are problem dependant. In adaptive control, information from the search process is used to alter the parameter values. With simple if-then rules based on the measures on the population the parameter values are constantly adapted to the search and the status of the population.

5.4.3.3. Convergence to Pareto-optimal Front and Diversity of Solutions

Multimodal functions have multiple optimum solutions many of which are local optimal solutions (Figure 5.12). Deb (1999) shows how a MOGA can get stuck at a local Pareto-optimal front if appropriate parameters of the algorithm are not used. A special case of multimodality is *deception* where at least two optima exist in the search space, but almost the entire search space "favours" the deceptive local optimum. Although the presence of multiple optima hinders the convergence to the true (global) Pareto-optimal front, a MOGA must at the same time be able to find the optimal or near optimal solutions as well as the local optima. Information contained in local optimal solutions is useful to DMs for choosing alternative optimal solutions as and when required. In practice, because of the spatial dependence of policy effectiveness, local constraints may make a country-wide optimum solution infeasible to implement.

In order to find and maintain local optimal solutions a MOGA can be modified to include an explicit diversity preserving operator. The mutation operator is often used because it can help find different optimal solution. However, it can help little in preserving local optimal solution over a large number of generations as it has a constructive as well as a destructive effect. Most of the popular diversity preservation techniques in use today are based on the concept of crowding. As the name suggests crowding of solutions anywhere in the search space is discouraged. This can be done by degrading the fitness of crowded (similar) solutions.

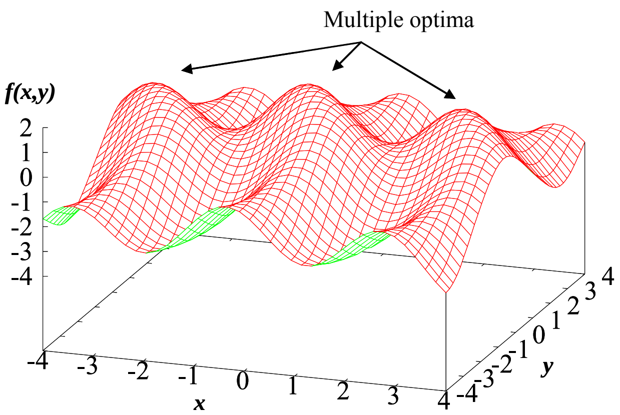

Figure 5.12. A simple multimodal function: *f(x,y)* = $\cos(2x+\sin(y)) + \cos(y) - 0.05(x^2 + y^2)$

However, there are some specific features, which if present in the global Pareto-optimal front, can cause a MOGA difficulty in maintaining diverse non-dominated solutions. For instance, non-convexity of the Pareto-optimal front may be a potential problem in GA implementations where the fitness of a solution is assigned proportionally to the number of solutions it dominates. This fitness assignment favours solutions intermediate in a non-convex Pareto-optimal Front, thereby causing an artificial bias towards the intermediate portion of the Pareto-optimal region (Deb, 1999). Another difficulty arises when the Pareto-optimal font is discontinuous. Although solutions within each disconnected sub-region may be found, competition among these solutions may lead to extension of some sub-regions thereby reducing solution diversity.

5.5. Summary

Multi-objective optimisation to be distinct from maximizing or minimising, uses the concept of Pareto optimality which involves balancing and harmonizing multiple criteria. There exist a number of classical and evolutionary methods for multi-objective optimisation the basic principles of some of their typical representative methods were described. All of the classical methods presented here, in order to handle multiple objectives convert a multi-objective optimisation problem into a single-objective one by assigning weights to each original objective functions. These weights should not be interpreted as the relative importance of objectives. However, there are at least two problems with this superfunctional aggregation of multiple objectives.

Firstly, criteria and attributes referring to different descriptive domains (*e.g.* environmental loading assessed in kilograms of chemicals per hectare vs. labour productivity expressed in added value per hour) reflect incommensurable qualities. (Benett *et al*, 2004) Scalarization requires that the values associated with specified objectives be standardised to ensure that the analysis remains unbiased by objective-specific units. It also requires a priori knowledge of the maximum and minimum obtainable objective values, and this knowledge is sometimes difficult to obtain. In other words, it is not possible to obtain a common measure of non-equivalent criteria; something that Munda (2004) calls *technical incommensurability*.

Secondly, even if a set (uniform-spaced in the absence of any additional knowledge) of weight vectors are used iteratively, it is likely, in sufficiently nonlinear problems, that the optimal solutions set will not be uniformly spaced (Deb, 2001). Therefore, finding a set of Pareto-optimal solutions first and then choosing the compromise solution, is considered by some authors (*e.g.* Deb, 2001; Zeleny, 2005) less subjective and hence a better practical strategy. This is the approach followed by MOGAs. The effort is made in finding the set of trade-off optimal solutions by considering all objectives equally important. After a set of such trade-off solutions are found, the user can then use higher-level qualitative considerations to make a decision.

Both classical and evolutionary methods have different strengths and shortcomings under different situations and choosing the most appropriate approach depends heavily on the characteristics and requirements of the underlying problem. The following chapter explains why MOGAs are better suited for addressing the multi-objective bi-level policy optimisation problem.

6. The Agricultural Policy Optimisation (APOLO) Model

6.1. Introduction

This chapter presents the difficulties associated with the agricultural policy optimisation problem and explains why it falls into the category of BLP. The focus of attention is on why MOGAs are suited to solve BLP problems and how they can be modified to work in conjunction with standard MP whereby multi-objective bi-level programming problem can be solved. More specifically, it shows how SASEM, which was presented in Chapter 4 can be nested in the NSGA-II algorithm, which was presented in the previous chapter, in such a way as Pareto-optimal or near Pareto-optimal solutions for the MOBLP problem of agricultural optimisation can be obtained. The integrated model is called APOLO (Agricultural Policy Optimiser). In the last section, the ability to solve single- and multi-objective BLP problems is demonstrated by using test problems.

6.2. Policy Optimisation and Bi-level Programming

Policy optimisation here refers to the process of finding of all possible combinations of policy measures the ones that best achieve the policy objectives and goals subject to available resources and various constraints. For example, when the policy aims at a multifunctional agriculture, where the set of goals includes food production, provision of environmental goods and services and rural development, one wants to know if any support payments or taxes are justified, where they should be directed to and at what levels. In a model the policy decision variables represent the policy instruments. As farmers respond to policy changes by altering their production patterns, the state of the agri-environment changes as well. For example, if beef prices decrease relative to cereals, farmers are likely to shift land use towards cereal production. One possible effect of these land management changes could be higher levels of soil loss and water quality degradation. Therefore, the prediction of the farmer's reaction to policy changes has a central role in agricultural policy analysis for it determines the production pattern of farming systems and ultimately the level of achievement of the multiple policy goals.

A distinguishing characteristic of the problem of agricultural policy optimisation is that policy makers' and farmers' objectives do not necessarily coincide; on the contrary they may be in conflict. In addition, the objectives of each decision-making unit may, in part, be determined by the actions of the other unit. To observe their goals policy makers may be able to influence the actions of farmers with their policies but not completely control their behaviour. From all the possible policies that are feasible from the policy makers' perspective, only those that are also satisfactory to the farmers are considered acceptable (can be implemented). Likewise, from all possible policies that are feasible to the farmers' problem, only some are approvable by the policy makers. The intersection, **A∩B**, of the policy-makers feasible space, **A**, and the farmers' feasible space, **B** forms the farmers' rational reaction region and is the set of feasible policy solutions. The optimal policy solution lies somewhere in the reaction region (Figure 6.1). The task for a policy optimisation model is to identify the Pareto-optimal set of policy solutions from within the intersection of solutions which are acceptable to policy-makers and farmers. To be able to handle this task the model should have a hierarchical optimisation structure.

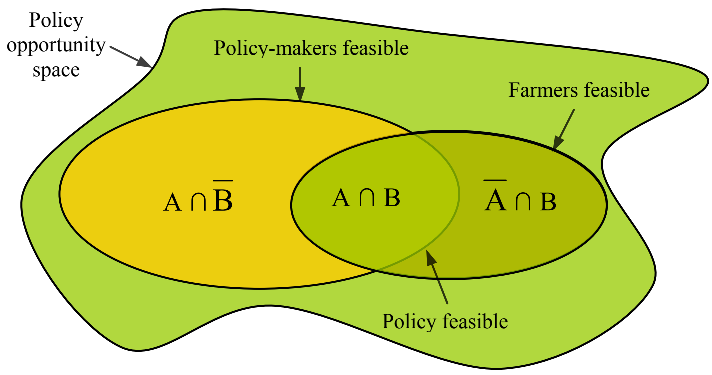

Figure 6.1. Schematic representation of the feasible space of policy solutions

Consider a policy MP model and a farmers MP model. Let, for example, SASEM to be the latter model. Its standard algebraic formulation is as follows:

Find the activity vector $\mathbf{x} = (x_1, x_2, \ldots, x_j)$

that maximises $$TGM^s = \sum_{j=1}^{J} (gm_j + sub_j) \cdot x_j - \sum_{j=1}^{J} C_j(x_j) \qquad (1a)$$

Subject to $$\sum_{j=1}^{J} a_{kj} x_j \leq b_k, \quad all\ k = 1\ to\ K \qquad (1b)$$

$$x_j \geq 0, \quad all\ j = 1\ to\ J \qquad (1c)$$

where $\boldsymbol{x}$ is the vector of production activities, x_j, equation (1a) is the farmers' objective function to be maximised namely, the total gross margin TGM^s, gm_j is the gross margin coefficient of x_j, sub_j denotes some policy variable, for instance , direct payments, $C_j(x_j)$ is the total variable costs of x_j. In the constraint equations (1b) a_{kj} is an input/output coefficient and b_k are relevant constraints, such as total available land etc.

Let now assume that the policy problem is to:

Find the policy variables vector $\mathbf{sub} = (sub_1, sub_2, \ldots, sub_j)$

such that

$$min\ PO = \sum_{j=1}^{J} z_j x_j \qquad (2a)$$

subject to $$sub_j \in [sub^{(L)}, sub^{(U)}] \qquad (2b)$$

where z_j gives the coefficient of an activity's contribution to the policy objective function.

The traditional way in which hierarchical decision problems of this sort are studied is to solve model (1a)-(1c) for a range of values of sub_j leaving it to the policy-maker to evaluate his/her objective function for the resulting activities x_j. This approach ensures that the policy solutions is feasible however, as Candler *et al* (1981) put it, it makes little sense, having gone to the expense of constructing and calibrating a farmers' model (such as SASEM), to fail to find the optimum feasible policy. Clearly it makes more sense to (re)formulate the two optimisation problems into a hierarchical bi-level programming problem where the Leader is the policy problem and the follower is the farmers' problem. The standard general[1] BLP problem is given by

$$\min_{\mathbf{x}} \quad F(\mathbf{x},\mathbf{y}) \tag{3a}$$

$$\text{subject to } \mathbf{G}(\mathbf{x},\mathbf{y}) \leq \mathbf{0} \tag{3b}$$

$$\min_{\mathbf{y}} \quad f(\mathbf{x},\mathbf{y}) \tag{3c}$$

$$\text{subject to } \mathbf{g}(\mathbf{x},\mathbf{y}) \leq \mathbf{0} \tag{3d}$$

where $\mathbf{x} \in X \subseteq \Re^n$, $\mathbf{y} \in X \subseteq \Re^m$ and $F, f : \Re^n \times \Re^m \to \Re^1$, $\boldsymbol{G} : \Re^n \times \Re^m \to \Re^p$, and $\boldsymbol{g}$: $\Re^n \times \Re^m \to \Re^q$ are assumed to be continuous, twice differentiable functions.

However, these types of problems are generally difficult to solve because evaluation of the upper-level objective function requires solution of the lower-level optimisation problem. Furthermore, because the lower-level problem is in effect a non-linear constraint, the problem is non-convex with more than one local optima (Yin 2000). The presence of multiple local optima in a problem restricts the use of standard linear and non-linear programming algorithms to solve it (Fortuny-Amat and McCarl, 1981; Onal *et al*, 1995; Martin *et al*, (2000). Omar and Blair (1990) show that solving BLP is NP-hard which makes it unlikely that there is a good, exact algorithm. This is true even for the simplest of forms where all the functions are linear and all the variables are continuous. Although exact approaches, such as algorithms based on Kuhn-Tucker optimality conditions of the follower's problem, or on variable elimination, handle the nonconvexity in different ways, their success has been limited (Bard, 1998 pp 361). A number of heuristics, (also known as global

[1] General in the sense that it includes all the versions i.e. linear BLP with continuous variables, linear BLP with discrete variables and, convex BLP. Depending on the functional forms of *F*, *f*, *G*, and *g* different versions of the BLP problem are obtained.

optimisation techniques) which have been proposed for solving more general classes of nonconvex problems (Horst and Tuy, 1995), can offer additional possibilities.

From the previous chapter it follows that the strength of genetic algorithms is their ability to handle multiple local optima, integer or discrete variables, as well as multi-objective problems. Specifically,

(1) GAs can work with a coding of the variable values instead of the values themselves; therefore, they can be easily adjusted to handle the integer or discrete variables;

(2) GAs use only objective function information, not derivatives or other auxiliary knowledge. Also, they use probabilistic transition rules, not deterministic rules. Therefore they can deal with the non-smooth, non-continuous and non-differentiable functions common in practical optimisation problems.

(3) Unlike most classical search and optimization problems, GAs search from a population of points, not a single point. Therefore GAs can conduct global search and thus are likely (and unique) candidates for finding multiple Pareto-optimal solutions simultaneously.

The major drawback of GAs however, is that, in problems that can be solved by traditional algorithms, they are significantly slower compared to the latter (Dorsey and Mayer, 1995). For example, practical MP models are most frequently simple LP or quadratic programming models with a large number of variables and constraints. Most linear and quadratic programming packages can solve these models quickly and accurately. It is unlikely that a heuristic method such as a GA would solve these problems any faster or with higher level of accuracy. Additionally, in cases where the follower's model has already been developed in a specific MP environment (*e.g.* GAMS, LINDO etc.) it makes little sense to set up again the same models in a GA. In order to solve the upper-level problem, the GA does not need to solve the lower-level models if their solution can be derived from elsewhere quicker and more efficiently; all that the GA needs is the solution information to navigate through the search space.

It follows that a "hybrid" MOGA algorithm which manages the two problems separately with two different specialised algorithms is more efficient and requires shorter execution times than a standard one. The next section shows how this approach is implemented in the APOLO model.

6.3. The Agricultural Policy Optimisation (APOLO) Model

APOLO implements the same "hybridisation" rationale as in Anandalingam *et al*, (1989), Yin (2000) and Yin (2002). The MOBLP problem is divided into two modules one for each optimisation level. The one called the Leader Module handles the upper level problem by using a MOGA, and the other called the Follower Module handles the model (such as SASEM) for the lower level problem by using a classical optimisation algorithm. The Leader Module is a simple modification of the original MOGA that allows for an interface with an external process, such as the Follower Module, whereby communication (i.e. exchange of data) between the two procedures is facilitated. The communication involves three stages: (1) for each possible policy solution, PS_p, (i.e. a set of policy decision variable values) considered by the GA, the MP model is solved by the embedded (Follower) module; (2) the solution of the MP model (i.e. the farmers decision variable values), FS_p, is passed on back to the GA; (3) the GA derives the policy objective function values associated with FS_p and uses these values to evaluate PS_p.

There are a number of public libraries of MOGA programmes and classic optimisation software. As discussed in Chapter 5, APOLO uses the *Elitist Non-dominated Sorting Genetic Algorithm* (NSGA-II) developed by Deb *et al*. (2000) as the platform for solving the policy model. In order to model the response of farmers to candidate policy solutions, SASEM is used. The development, main features and validation of SASEM were described in Chapter 4. Based on the arguments and findings presented there SASEM is a suitable tool for modelling the aggregate response of farmers in Scotland to policy and market changes. The Solver Dynamic Link Library (*Solver DLL*) is an optimisation software package developed by Frontline Systems Inc. It is used as the platform for setting up and solving the SASEM model. The original NSGA-II programme is implemented in C therefore, to greatly simplify the integration, C was also chosen as the implementation environment for SASEM. The original code of NSGA-II was modified slightly in order for an interface between the GA and SASEM to be created.

Figure 6.2 shows the flowchart of the various stages[2] of the hybrid algorithm. Analytically the stages are:

Stage 1. Define Policy Objectives and Variables

At this stage the set of policy objective functions and the set of policy variables are chosen based on the criteria and policy instruments associated with the aims and scope of the policy experiment. A set of policy variable values makes up a policy solution. Objectives that are dependant on the farmers' response should be expressed as functions of decision variables of SASEM.

Stage 2. Initialise Population of Solutions and Counter

At this stage an initial population of *P* policy solutions, PS_p, is created randomly to serve as the starting point for the algorithm. To achieve a pseudo-chromosomal representation of the solutions in the algorithm, the binary alphabet {0, 1} is used for encoding their values in a string of 0s and 1s. A counter called *gen* (for generation) is set to 1.

Stage 3. Parameterise and Solve the SASEM Model

For each policy solution, PS_p, in the population the GA calls the Follower Module. Every time the Follower Module is called, PS_p (i.e. the vector of policy decision variable values, $\mathbf{x}_p^{\mathrm{P}}$) is passed into SASEM as a vector of fixed parameters. The parameterised SASEM is then solved using the *Solver DLL* and the optimal solution,

[2] It should be made clear that all the stages of the hybrid algorithm (APOLO) apart from stage 3 and 4 are those of NSGA-II; stage 3 is added and stage 4 is slightly modified.

FS_p, (i.e. the vector of farmers' decision variable values, $\mathbf{x}_p^f$) is obtained. After SASEM is solved, the Follower Module returns the solution, FS_p, to the GA.

Stage 4. Measure Performance of Policy Solution

The overall performance of each policy solution, PS_p, with respect to all the objectives is measured by the levels of achievement of these objectives. These are obtained by substituting the leader and follower decision making variables with PS_p, and FS_p respectively in the objective functions. At the end of this stage every policy solution is associated with a set of objective function values. These are used by the GA in the next stage.

Stage 5. Assign Fitness to Policy Solutions

Fitness evaluation involves assigning a fitness value to each policy solution, PS_p, in the current population. Because individual solutions are associated with a set of objective function values they can be compared on the basis of whether one solution dominates the other or not. Then, the population is sorted into different non-domination levels by applying the non-dominated sorting procedure described in section 5.4.1.4. In short: non-domination level 1 includes the best non-dominated solutions of all population; in order to find the solutions of the n^{th} level, all solutions up to the n-1 level are temporarily disregarded and the non-dominated solutions of the remaining population are then found. After the non-dominated sorting is complete each chromosome is assigned a fitness value equal to its non-domination level (thus the fittest solutions have the lowest scores).

Stage 6a. Selection and Elite-preserving Operator

The algorithm performs *crowded tournaments* (see section 5.4.2.1) to select from the current population of policy solutions the parents (mating pool) for the next generation. The same solution can be represented in the mating pool for as many times as the number of tournaments it has won. Thus, good solutions have more copies in the mating pool than less good solutions. The elite-preserving operator makes sure that the fitness of the population-best solution does not deteriorate. In this way, good solutions found early on in the run are not lost unless better solutions are discovered.

Stage 6b. Reproduction or Crossover Operator

Once the mating pool (parents) has been formed, the crossover operator acts to produce offspring by randomly selecting a pair of parents and replacing part (single-point crossover) or parts (multi-point crossover) of the string of one parent with the corresponding part of the string of the other parent.

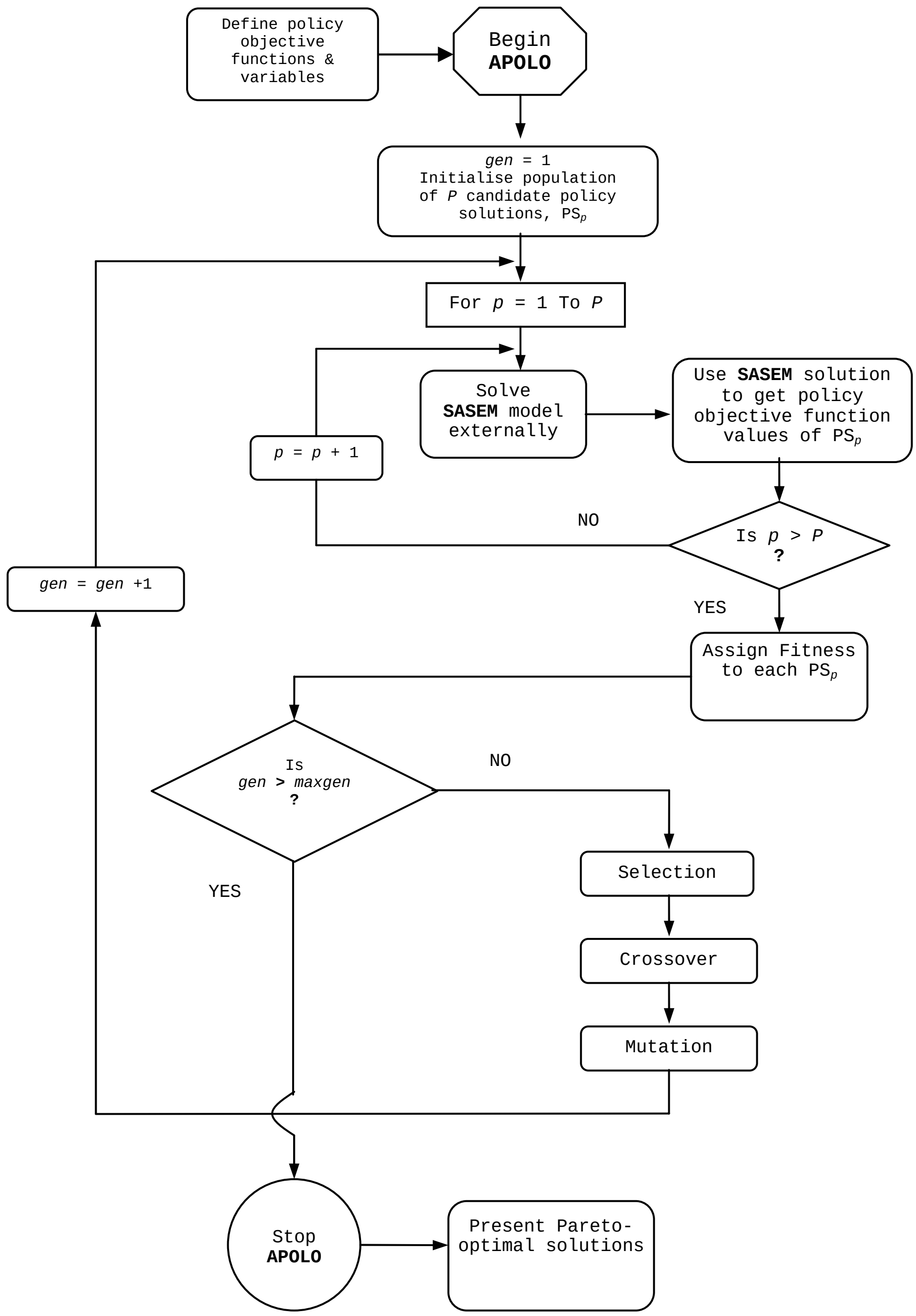

Figure 6.2**.** A flowchart of the stages of APOLO

Stage 6c. Mutation Operator

The mutation operator introduces some random alteration of the genes, e.g. 0 becomes 1 and vice versa to enhance diversity.[3]

Stage 7. Increase generation counter

The generation counter is incremented by 1 and the procedure is repeated for the member solutions of the new generation until a maximum number of generations, *maxgen,* (which is problem dependant) have elapsed. At the end of one cycle a Pareto-optimal set is produced from which trade-offs between the objectives can be quantified and a compromise solution is possible to be reached.

6.4. Validation of APOLO

The effectiveness of the model is illustrated by using a test problem obtained from the literature (Martin *et al*., 2000). Also a variant with two objective functions for the Leader is examined. The test problem is very simple (linear with two variables) nonetheless, it represents a number of features that create difficulties for BLP optimisation algorithms. The algebraic formulations of the standard version and its variant are presented in Table 6.1. Despite the fact that the upper and lower level objective functions and all constraints are linear, and that each level's problem is strictly concave, this problem contains multiple optima. As Figure 6.3 illustrates, the

[3] In the context of APOLO the action of the mutation operator represents arbitrary and uncoordinated (with respect to other policy instruments) switches in policy. Typically, such changes are unlikely (although it might be argued that they are more likely than their genetic analogues in biological systems so the mutation rate used is of the order of about one bit changed per thousand.

overall feasible space **S** is a convex polyhedron however, the actual problem is non-convex.

The vertical axis represents variable y whereas the horizontal axis represents variable x. The slope and direction of the objective functions F and L_1 are shown by the dashed lines and arrows respectively. The Leader controls variable x and for any fixed choice of x, the Follower chooses the value of y which maximises objective function F over the feasible space, which is a convex set. This results in the rational reaction space, **S′**, which is depicted as the heavy dashed line. The Leader will obviously choose a value for x which maximises objective function L_1 over **S′**. A problem however arises due to the existence of two local optima at points ‘B’ and ‘D’. Any movement away from these points (on the rational reaction space) increases y and therefore, reduces the Leader’s objective value.

Table 6.1. The BLP test problem and its variants

Test Problem 1	
Standard	Variant
Find x^L and y^F that max L_1 = - y (Leader) Subject to max F = - x + 2 y (Follower) Subject to $-x - 0.5y \leq -2$ $-0.25x + y \leq 2$ $x + 0.5y \leq 8$ $x - 2y \leq 4$ $x, y \geq 0$	Find x^L and y^F that max L_1 = - y (Leader) max L_2 = - x (Leader) Subject to max F = - x + 2 y (Follower) Subject to $-x - 0.5y \leq -2$ $-0.25x + y \leq 2$ $x + 0.5y \leq 8$ $x - 2y \leq 4$ $x, y \geq 0$

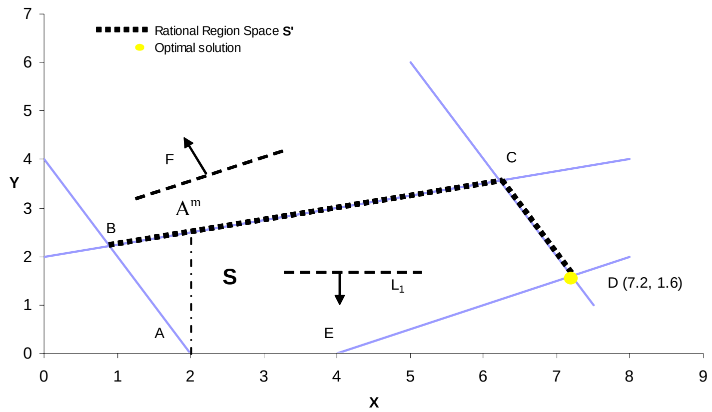

Figure 6.3. Diagrammatic representation of the BLP test problem (standard)

The globally optimal solution for the Leader is observed to be at point 'D' i.e. to set $x = 7.2$ in anticipation that the Follower will set $y = 1.6$. This solution however is not Pareto optimal. Point 'A' is Pareto superior because it provides higher value (utility) for both the Leader and the Follower[4]. Marcotte and Savard (1991) have pointed out that even in the simplest case of linear BLP the globally optimal solution will not, in general, be also Pareto optimal.

The presence of more than one local optimum in the test problem restricts the use of standard linear and nonlinear algorithms to solve it. An alternative formulation for relatively small problems is to replace the follower level problem with its Kuhn-Tucker optimality conditions and use them as constraints to the leader level problem. Optimisation results in a feasible and locally optimal solution; however, it cannot guarantee a convergence to the global optimum. When it was attempted to solve the standard test problem with the equivalent Kuhn-Tucker formulation in GAMS-MINOS the globally optimal solution was not uncovered. APOLO was tested to see if it is able to converge to the globally optimal solution.

For the standard version of the problem the parameter setting for the GA was as follows: number of generations = 100; probability of crossover = 0.9; probability of mutation = 0.08; three different population sizes, P = 10, 20, 30; two different numbers of bits assigned to GA's variable x, *bits* = 10, 20 with lower limit 0 and two upper limits 10 and 50. These settings gave twelve runs in total: four for each population size, of which two were for each number of bits of which one was for each upper limit. The model was run on an Intel Pentium 2.0 GHz personal computer, with total CPU time between 15 – 35 sec depending on the population and generation numbers. Figure 6.4 shows the results from each run. They clearly

[4] However, it can be argued that at point 'D' the Follower does not forgo any utility, his utility is at the maximum possible level at this point, whereas if he compromises with point 'A' instead of choosing the maximum possible (point 'A^m') it means that has to forgo some of his potential utility. Taken this opportunity cost into account one might suggest that globally optimal point 'D' is also Pareto-optimal.

indicate that APOLO has no problem to converge to a near global optimum for most of the parameter settings.

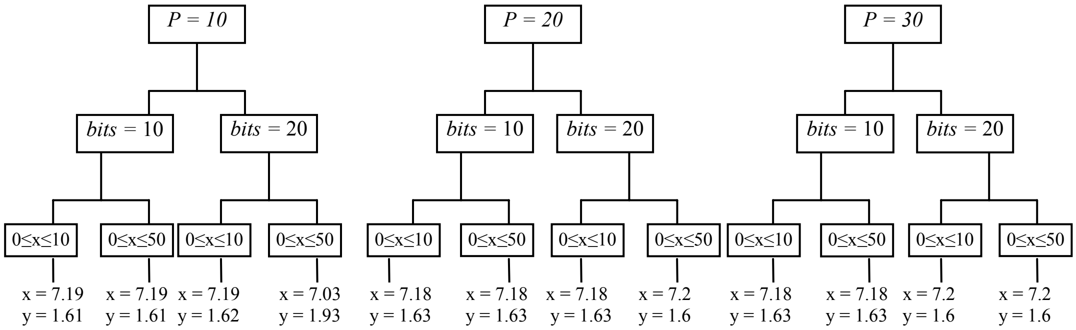

Figure 6.4**.** Results from the APOLO model for the test problem (standard version)

Since the problem of policy optimisation besides being bi-level is at the same time a multi-objective one, problem 1 was modified slightly for the Leader to include a second objective function. The test for the model is to find the Pareto-optimal set in the rational region space **S**. The model was run a number of times with different parameter settings and it consistently generated the set which is depicted as the heavy red dots in Figure 6.5. This set represents the trade-off curve between Leader's two objectives. It can be seen that the Pareto-optimal set is discontinuous. Solution point 'B' dominates all feasible solutions (on the rational region space) that are above the red fine dashed line whereas, all other feasible solutions under the red line are Pareto-optimal solutions; *i.e.* no one solution dominates the other. By looking at the graph one can confirm that the Pareto-optimal solutions found by APOLO are also the true Pareto-optimal solutions. Thus, the result of this test demonstrates the model's ability to solve multi-objective bi-level linear programming problems.

To test the model's ability to converge to globally optimal solutions in problems with quadratic objective functions the problem shown in Table 6.2 was used. The model was successful, under various GA parameter settings, to recover the global optimum using 15-25 sec of CPU time.

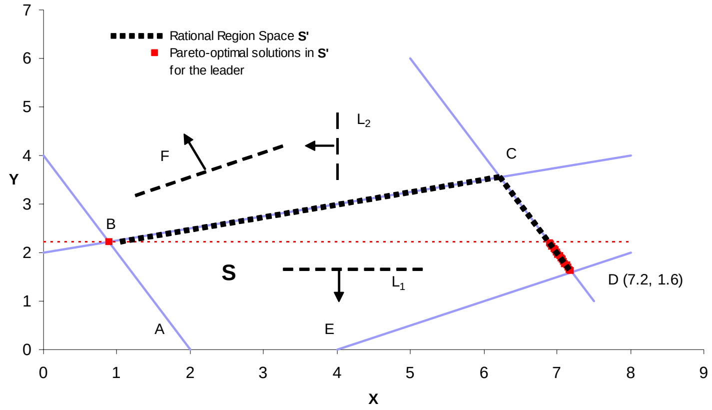

Figure 6.5. Pareto-optimal solutions found by APOLO for the Bi-objective BLP problem (variant)

Table 6.2. A convex quadratic – quadratic BLP problem

Test Problem 2
Find x^L and y^F that
min $L_1 = (x-5)^2 + (2y+1)^2$ (Leader)
Subject to
min $F = -1.5x - (y-1)^2$ (Follower)
Subject to
$-3x + y \leq -3$
$x - 0.5y \leq 4$
$x + y \leq 7$
$x, y \geq 0$

6.5. Summary

The multi-objective and hierarchical nature of the agricultural policy optimisation problem suggests that it should be formulated as a BLP problem (3a-3d)). All widely-used NLP solvers assume that all problem functions possess continuous first derivatives with respect to the decision variables (Bard, 1998). Unfortunately, the vector **x**(**y**) which is returned form the solution of the follower's problem (3c-3d), is not necessarily differentiable everywhere or even continuous in **y**. This makes it difficult to apply nonlinear programming theory to the BLP directly (Bard, 1998). The absence of global exact algorithms for BLP problems necessitated the use and modification of a heuristic optimisation technique. This involved the integration of a multi-objective genetic algorithm with standard mathematical programming. This approach has been implemented to build APOLO, a model for solving agricultural policy problems. Although based on very simple test problems the results provided here give some evidence that APOLO can find the global optimum among multiple optima and the Pareto-optimal solutions for the Leader's multiple objectives in hierarchical optimisation problems with two levels. Of course, further investigation with larger and more difficult problems is needed before one can assert with certainty that the model is robust with all types of BLP problems. However, an appealing feature of APOLO is that it can readily work with MP models that have been developed in different programming environments, suggesting that their transition to BLP models is a simple matter.

PART III

Multi-objective Bi-level Optimisation of Agricultural Policy: An Application to Scottish Agriculture

7. Optimisation of Scottish Agriculture Policy: Policy Design

7.1. Introduction

A first step before applying the APOLO model in various hypothetical situations of agricultural policy in Scotland, is the development of relevant policy scenarios. The primary aim of this chapter is to demonstrate how the scenarios considered for analysis were designed. To formulate relevant and realistic policy scenarios, it is important to take under consideration the political and socio-economic landscape upon which the farming systems in Scotland currently operate. International agreements on trade liberalisation and environmental along with other non-trade considerations of agricultural production practices are two of the principal factors shaping this landscape.

Trade liberalisation of agricultural commodities has led to adjustments in production and consumption patterns and these changes may affect the environment and the structure of rural economies and societies. Neo-classical economists (e.g. Anderson 2000; Harvey 2003) often advocate trade liberalisation arguing that removal of production-tied support can result in more efficient (and presumably less intensive) farming methods whereby environmental quality would improve as a lateral benefit. In contrast, critics of the agricultural trade liberalisation alleged that predominant reliance on market signals would have the opposite effects. The loss of ecological services provided by farmlands due to the abandonment of crop and livestock production has been a frequently cited risk. For example, the loss of open agricultural landscapes and the loss of species dependant upon these semi-natural ecotypes have been recognised by some authors (e.g. Sumelious, 1997). These two opposite hypotheses appear competing, but the outcomes of trade liberalisation are more complex and perhaps not mutually exclusive (Prestegard, 2003). There can be diverse environmental, as well as socio-economic, effects given the different trading regimes, agricultural systems, natural environments, agri-environmental and rural development programmes across countries.

Non-trade concerns that encompass food security, environment, structural adjustment, rural development, poverty alleviation and so forth are often reflected on the concept of multifunctionality (WTO, 2001). Multifunctionality refers to the idea/concept that an economic activity may have multiple outputs and may contribute to several societal objectives at once. While the term has been used by the European Union in international negotiations of agricultural trade it was first recognised at international level in the Rio declaration on sustainable development in 1992 and later by the FAO at its World Summit in 1996. In March 1998, the OECD agricultural ministers Committee (OECD, 2001) defined the multifunctionality of agriculture as follows:

> "…beyond its primary function of supplying food and fibre, agricultural activity can also shape the landscape, provide environmental benefits such as land conservation, the sustainable management of renewable natural resources and the preservation of biodiversity, and contribute to the socio-economic viability of many rural areas".

Despite consensus about the existence of multiple benefits of agriculture to society, there is a controversy about which instruments to implement in order to promote multifunctionality. Some authors argue that under certain conditions production tied support of agricultural commodities should be justified (Vatn, 1999; Prestegard, 2003), others suggest that providers of public good services that are positively valued by society should be compensated according to the Provider Gets principle (e.g., Hodge, 2000), while other authors strictly oppose any modifications of the Green box criteria on the grounds of multifuncionality (e.g., Anderson, 2000; Harvey, 2003). They argue that policies promoting joint production of agricultural and environmental outputs by internalising domestic externalities may affect quantities produced, trade flows and world prices and may impose burdens on trading partners. Under the WTO negotiations the legitimacy of policy interventions are assessed in terms of to what extent they distort the markets.

According to Prestegard (2003) even if there are some possibilities for transforming externalities and public goods into marketable commodities, and that voluntary organisations may play a certain role in providing positive externalities/public goods, there will still exist a significant role for governmental support to enhance agricultural multifunctionality in OECD countries. Harvey (2003) recognises a twofold focus of sustainable policies in first, getting the price of agricultural marketable output “right”, and second, properly reflecting the public or social values of the non tradable goods back to landowners and land stewards. Overall, there seems to be consensus that a holistic, integrated agricultural strategy has to move away from market price support to a system which better balances and integrates the role of farmers as competitive providers of agricultural commodities on the one hand and, providers of public goods on the other.

To this end, significant changes have taken place in recent years in the types of policy measurements used for delivering support to farmers. The policy mix was totally dominated by price-related measures in the 1970’s and 1980’s. It has been gradually expanded to include new types of policy instruments, many of them in the form of budgetary support, paid directly to farmers. These new instruments are conditional on a complex range of implementation conditions and criteria including standards (good agricultural and environmental conditions for the UK), quotas and management agreements targeted at regions or individual farms. In the European Union for example, the key element of the 2003 CAP reform is the introduction of a single payment scheme which replaced a multitude of premiums formerly coupled to output (crop area & livestock numbers). Direct payments of the rural development program are not affected by this reform.

From the above it follows that agricultural policy should be designed in a way which ensures that the social goals of multifunctionality and sustainability are achieved in a minimally trade-distorting way.

7.2. Policy Design

In the context of the policy analysis presented in the application, a certain policy design comprises a certain policy scenario. Each policy scenario consists of two elements: Policy objectives and policy instruments for achieving the objectives. It is intuitive that different choices of these two elements and their various combinations can lead to a considerable number of alternative scenarios. This section presents only the set of policy objectives and instruments that were considered the most relevant and at the same time the APOLO model was able to handle.

7.2.1. Policy Objectives

European and consequently Scottish agriculture is suffering from an internal conflict of interests. There is a desire and, to a certain extent, a commitment to the sustainable development and the multifunctional character of agriculture, but at the same time there is a promotion of an increase in production in the primary objectives of the CAP (Nilsson, 2004). Specifically, for example, one of the original objectives of the CAP set down in the 1957 Treaty of Rome, which has remained in place since then, is "to increase productivity by promoting technical progress and ensuring the rational development of agricultural production" (Treaty of Rome, 1957).

Recalling that within APOLO, in order to evaluate the fitness of any policy solution, policy objective functions must be defined and formulated. Hence, a set of quantifiable attributes related to the policy objectives has to be identified first. The choice of the appropriate attributes is problem dependent. In an agricultural policy context they should be highly correlated with agricultural sustainability.

The literature suggests several approaches for evaluating sustainable agroecosystems. Soule and Piper (1992) for example, describe the health of an agroecosystem in terms of the characteristics such a system exhibits, and characterise the indicators of an agroecosystem in terms of its structure and function. Lancaster's (Lancaster, 1966) theory of consumer demand on the other hand, asserts that humans demand the characteristics inherent in goods and not the goods *per se*. As Yiridoe and Weersink (1997) argue, this concept may be applied to the notion of agroecosystem sustainability, suggesting that humans demand the characteristics associated with a healthy agroecosystem rather than merely an agroecosystem that is sustainable. It is these characteristics which are reflected and quantified in sets of indicators of agroecosystem sustainability (OECD, 1997; European Commission, 2000; MAFF, 2000). Such indicators, therefore, consist of the information pool (as the various criteria) that the evaluation of policy fitness could be based on, and their values reflect the levels of achievement of the policy objectives.

Lara and Stancu-Minasian (1999) have argued that with appropriately chosen objectives of sustainability, the partitioning of the set of potential policy actions (solutions) between *efficient* (Pareto-optimal) and *non-efficient* (dominated) is the same as classifying those actions as either more sustainable or less sustainable. Therefore by adjusting the definition of Pareto optimality appropriately, any policy solution can be defined as more sustainable if there is not another alternative solution that achieves at least the same or better performance with respect to every desired objective, and strictly better with respect to at least one objective.

As explained in the preceding paragraphs, to facilitate an adequate projection which is able to represent the most important features of the state of fitness of agriculture a set of indicators should be used. Piorr (2003) provides a list of demands on indicators used in decision making. Among others, they should have at least some of the following policy related attributes: 1) provide a representative picture of environmental, agricultural and rural conditions, 2) be simple to use and easy to interpret for different users, 3) provide a basis for regional, national and international comparisons, 4) be either national in scope or applicable to regional issues of national significance, and 5) assist individual decision-makers of the private sector as well as trade and industry. In addition, the choice of indicators is subject to the particular aims and focus of the analysis, as well as on practical issues such as data availability.

Given the limitations in availability of data and biophysical models of the relationship between production activities and indicators, as well as the need to keep the model manageable, the number of indicators considered in this study was limited to a maximum of five. If the general aim of the policy is to enhance a sustainable and multifunctional agriculture (in contrast, for example, with a policy aimed specifically to improve farmers' income) the policy objectives should reflect relevant indicators of sustainability and multifunctionality. It is assumed that it is possible to represent adequately the socioeconomic and the ecological state of agricultural systems, and consequently the performance of the policy that results in such a state, on the basis of three socio-economic and two biophysical indicators respectively. These indicators and the corresponding objective functions are described in the following paragraphs.

7.2.1.1. Land and Labour Productivity

Giampietro (2004) used the land and labour productivity indicators to show the effect of technical progress in agriculture and changes in the use of technical inputs on terrestrial ecosystems. Increasing labour productivity is related to the need of improving the standards of living of farmers and making more labour available for other economic sectors. Increasing the land productivity is related to the need of covering the food requirements of a growing population and increasing the per hectare returns. Release of labour presently associated with uncompetitive farming for other more productive and socially desirable uses contributes also to rural development according to Harvey (2003). Since the physical productivity of the various production activities is measured in different units, (*e.g.* tonnes of wheat produced per unit of labour or tones of milk produced per unit of land), the total productivity of agriculture, (*i.e.* the productivity of all the activities combined), can only be summed up if the various unit are converted into a common one. Hence, after all output is converted into monetary value land productivity can be defined as the total gross margin generated per total land used, and labour productivity as the total gross margin generated per total hours of labour used, measured in £/ha and £/hr respectively. These can be expressed mathematically as follows:

$$LanPr = \sum_{j=1}^{J} gm_j X_j / \sum_{j=1}^{J} la_j X_j \qquad (7.1)$$

$$LabPr = \sum_{j=1}^{J} gm_j X_j / \sum_{j=1}^{J} lb_j X_j \qquad (7.2)$$

la_j is the coefficient that expresses the units of land used per unit of activity j, and lb_j is the coefficient that expresses the units of employment required per unit of activity j. Coefficient la is 1 for all cropping activities and 0 for input and livestock activities. The data on labour requirements are taken from the Farm Management Pocketbook (Nix, 2004). X_j gives the level of activity j and gm_j its gross margin. Equations 7.1 and 7.2 indicate that land and labour productivities are calculated excluding any policy payments per area or head; *i.e.* using only the gross margins of the activities.

7.2.1.2. Cost and Revenue of policy

In reality, the cost of policy intervention and the limits on the administrative capacities of government may form significant constraints on policy-making and implementation. In this study only the expenditure for all the payments and the revenue from all taxation were included in the model; any administrative costs or constraints were left out of the analysis. Rather than restricting the solution space due to a budget constraint it was preferred to treat the cost and revenue associated with a policy solution as objectives. In this way more policy options would be available for consideration. In addition, some inferences about modulation options could be made by comparing the expenditure of policy alternatives with the maximum allowable budget. If S_j and T_j denote the payment and the tax ratio per unit of activity j respectively, the two objective functions can be expressed as following:

$$PolC = \sum_{j=1}^{J} S_j \cdot X_j \tag{7.3}$$

$$PolR = \sum_{j=1}^{J} gm_j \cdot (1 + T_j) \cdot X_j \tag{7.4}$$

7.2.1.3. Nitrogen Pollution

The nature of inorganic nitrogen impacts on the agri-environment following field application depends on its loss through denitrification, ammonia volatilisation and leaching, which in turn depend on a number of factors, such as land use, soil cover and structure, local topography and climate, and species presence (Di and Cameron 2002). Reducing the total volume of inorganic N input may not be the most efficient means of securing reductions in risks to health and the environment (given an imperfect correlation between volume applied and environmental burden) nonetheless, input volume is a relatively low transactions-cost proxy for environmental burden, and hence a good indicator at least for policy design (Falconer and Hodge 2000). Suppose that n_j gives the coefficient that expresses the amount of nitrogen fertiliser in tonnes required per unit of cropping activity $j \in [1, J_1]$ then, the mean (per ha) usage of N fertiliser for the total area under agricultural production can be calculated by the following equation:

$$MNU = \sum_{j=1}^{J_1} n_j X_j / \sum_{j=1}^{J_1} X_j \tag{7.5}$$

Data on N fertiliser applications per ha for the cropping activities J_1 included in the model represent typical practices and are taken from the Scottish Agricultural College Farm Management Handbook (Chadwick, 2005).

7.2.1.4. Soil Loss

With regard to soil quality, the OECD working groups (OECD, 1999) considered that the most important impacts from agricultural activity on soil quality relate to soil erosion, organic matter loss, and loss of soil biodiversity. From those three the latter only is considered as a soil quality indicator in the present study. As regards organic matter loss, this cannot satisfactorily be predicted from cropping and management practices, so both OECD (1999) and MAFF (2000) rely on *ex-post* measurements to track these changes, making the corresponding indicator inappropriate for an *ex ante* model. On the other hand, an empirical model for predicting soil erosion, the Universal Soil Loss Equation (USLE) has been developed in the United States by Wischmeier and Smith (1979). The basic format of the equation is as follows:

$$E = R \cdot K \cdot L \cdot S \cdot C \cdot P \qquad (7.6)$$

Where E is the mean annual soil loss in t per ha, R is the rainfall erosivity factor, K is the soil erodibility factor in t per ha, L is the slope length factor, S is the slope steepness factor, C is the crop management and P is the erosion-control practice factor. The crop management factor represents the ratio of soil loss under a given crop to that from bare soil. Since soil loss varies with erosivity and the morphology of the plant cover, it is theoretically necessary to take account of changes in these during the year. However, values of C have only been derived in detail for the US and so Morgan (1995) recommends that for other countries it is more appropriate to use average annual values. These are presented in Table 7.1. Forage crops (legumes and grasses) provide the best soil protection because of their relatively dense cover. Small grains, such as wheat and oats, provide intermediate cover and offer considerable obstruction to surface. Bare soil obviously offers no protection.

Table 7.1. Mean annual values for C for different types of land cover

Land Cover	Average Annual Value of *C*
Bare soil	1.00
Grassland	0.02
Wheat	0.25
Maize	0.40

When all the parameters in USLE (7.6) are known the mean annual soil loss in t per ha for a given cropping activity, E_j, is also known and the total annual soil loss in t can be calculated from the following equation:

$$TE = \sum_j E_j X_j \qquad (7.7)$$

However, in this study the model is not spatial and therefore it is not known where each crop would be allocated, only its total area. Consequently, relevant data on site specific conditions cannot be used here, and E_j cannot be calculated. Nonetheless, a simpler approach can be followed to indicate susceptibility to soil erosion, where only the crop management factors C_j are needed. In the PMP model, since there are not different classes of land, any activity competes with the other activities for the same land. It follows that if the same erosion-control factor, P, applies to all crop activities, for any two different cropping activities competing for the same ha of land, the mean annual soil loss differs only in the crop management factor, namely the plant cover. Thus, the sub product $R_j \cdot K_j \cdot L_j \cdot S_j \cdot P_j$ in equation 7.6 can be treated as a common constant coefficient denoted $z_j = z$. Substituting $R_j \cdot K_j \cdot L_j \cdot S_j \cdot P_j$ with z and factoring it out, equation 7.7 now reads:

$$TE = \sum_j z_j C_j X_j = \sum_j z C_j X_j = z \sum_j C_j X_j \qquad (7.8)$$

By rearranging 7.8 becomes:

$$\sum_j C_j X_j = \frac{TE}{z_j} = TE_c \qquad (7.9)$$

TE_c can be defined as the total annual soil erosion susceptibility due to crop management. The mean (per ha) annual soil erosion susceptibility due to crop management defined as

$$ME_c = \sum_{j=1}^{J_1} C_j X_j / \sum_{j=1}^{J_1} X_j \qquad (7.10)$$

is used as an indicator of soil quality in the PMP model.

Summarising the preceding descriptions, the five possible policy objectives to be considered are: 1) maximise land productivity, 2) maximise labour productivity 3) minimise policy cost, 4) minimise mean nitrogen fertiliser use, and 5) minimise mean soil erosion susceptibility due to land cover. It should be noted again here that payments made to the production activities and taxes levied to inputs are not included when calculating marginal return (net of payments and/or taxes). To do otherwise would under-represent the trade off between environmental quality and agricultural productivity because higher payments would result also in higher productivity. Given this formulation of the problem, we expect that as public expenditure in the policy alternative design increases, environmental damage will decline and the productivity of land and labour associated with the policy will decline as land is allocated to less damaging but also less productive cropping activities.

Table 7.2. Policy objectives

	Indicators	Notation in the Objective Function / optimisation type	Code used in scenarios
Socio-economic	1. Land Productivity (£/ha)	*LanPr / max*	SO1
	2. Labour Productivity (£/hr)	*LabPr / max*	SO2
	3. Cost or Revenue of Policy (£ M)	*PolC, PolR / min, max*	SO3
Biophysical	1. Mean Synthetic Nitrogen Fertilizer Use (t/ha)	*MNU / min*	BO1
	2. Mean Annual Soil Erosion Susceptibility due to Crop Management (t/ha)	ME_c */ min*	BO2

7.2.2. Policy Instruments

A specific policy measure is effective if the intended policy objective is achieved and efficient if the objective is achieved at least cost (Flury *et al*, 2005). Prestegard (2003) argues that to achieve several objectives simultaneously at least as many policy instruments as there are objectives are needed. Of course, any given policy measure can have both positive and negative effects on more than one objective.

Also, the various policy instruments will influence each other, thus it will often be useful to set up (and evaluate) packages of policy instruments to obtain efficient policies that perform optimally with respect to all the policy objectives under consideration (Romstad *et al*, 2000), while at the same time minimising any trade-distorting effects. Prestegard (2003) points out three main challenges: 1) how to formulate the policy instruments, 2) how to determine their levels (*e.g.* the level of payments, taxes or quotas), and 3) which policy instrument should be accepted as minimally trade-distorting within the World Trade Organisation. While the last point is a major issue within the WTO, Burell (2001, pp 13) reminds us that in the presence of an unmarketable externality the market is already distorted. She continues by noting that direct payments aiming at correcting for market failure are hardly expected to have no effect on production levels and trade. This can easily mean that the definition of *minimally distorting* policy depends on what exactly the distortion is measured against. This is most often a presumed baseline situation referring to market supply and demand functions that can only be discovered econometrically. However, if the policy induced change in production is measured against the current (distorted) level, Burrell questions why should it be zero or minimal? (Burrell, 2001).

The agricultural policy instruments with potential distorting effects on agricultural trade include *regulatory instruments* (such as product or process restrictions or bans, technical regulations, resource use quotas), *economic instruments* (such as environmental taxes and charges, environmental subsidies, deposit-refund systems) and *voluntary schemes* (such as eco- labelling schemes) (Opschoor and Vos 1989, Office of Technology Assessment 1995).

Regulatory instruments try to influence the environmental performance of farms directly, by regulating the production processes or by abandoning or limiting the use of potential pollutants through licensing, standards and zoning. Regulatory action may restrict the availability of environmentally hazardous agricultural inputs, or prohibit the use of environmentally damaging production practices. It can be applied uniformly to all farmers, or may target specific farming operations or particularly vulnerable, or ecologically significant, agricultural areas. Recently, it has been combined with cross-compliance criteria which mean that a farmer's production practices must meet certain requirements in order to be eligible for monetary support. Farmers claiming support under one programme have to meet the conditions for that programme and certain obligations of other programmes. This makes a link across programmes, hence the term cross-compliance (Baldock and Michell 1995).

There is now widespread interest in the use of economic instruments to complement regulatory instruments for environmental management. Instruments of agri-environmental policy are labelled economic if they affect probable costs and benefits of alternatives, in this way influencing the decisions taken. Economic instruments are based on the polluter-pays and provider-gets principles and include mainly two categories: subsidies and taxes. A subsidy scheme might pay farmers who use environmentally friendly production practices such as integrated pest management or organic farming. When the issue is not one of pollution, but rather one of maintaining or enhancing the ecological services of agriculture, incentive schemes are particularly effective.

Taxation instruments include taxes on farm inputs which are sources of pollution, on farm emissions, or taxing farmers for their failure to meet required levels of environmental quality. An approach being tried in several countries is the use of input taxes to reduce the use of agricultural chemicals (see for example Falconer and Hodge 2000). Fertilizer charges as high as 100% are needed to reduce pollution significantly. However, taxes of only 10 - 20% may have a favourable effect. Kim and Kim (2000) reported that a tax of 100% on the nitrogen in chemical fertilizer leads to a reduction of 14.6% in fertilizer use and a fall of 0.3% in rice yield and 3.1% in farm income.

It is not known *a priori* which type of policy instruments are the most effective or whether some combination of different types would result in better achievement of the policy objectives. Hence, various types are tested to see if any gives superior solutions. In total five sets of policy instruments were selected for the analysis: three sets of economic instruments (single payment, land use subsidies and nitrogen taxation), one set of regulatory instruments (constraints), and one set comprising a mixture of the latter three.

7.2.2.1. Single Payment

In the European Union level and particularly in Scotland support mechanisms coupled to agricultural production have been replaced by a single decoupled payment subject to good agricultural practices and environmental standards. This new scheme called Single Farm Payment (SFP) is completely decoupled from production related activities. The payment is farm-specific and is based on the average of the payments the farm received over a three-year reference period from 2001 to 2003. It is an area-based direct payment instrument but with a much reduced likelihood of affecting the choice among alternative activities (OECD, 2006). However, according to Chau and DeGorter (2005) it can influence farm level output by covering losses from farming and thus preventing (inefficient) farmers from exiting the market. Nevertheless, the impact on total sectoral production is ambiguous since if a farmer exits the market land will not necessarily be idle, but can be taken over by more efficient farmers.

7.2.2.2. Unconditional (decoupled) Single Payment

This is the case where irrespective of what he or she produces the farmer will receive the payment. In principle, its effect on production would be limited to farm household resource allocation decisions. In its current formulation the SASEM model cannot simulate any effects on land use decisions caused by an unconditional single payment. As far as the SASEM model is concerned, this is equivalent to zero coupled payments and consequently activity levels in the model's solution depend only on market-dependent prices (gross margins) of outputs and inputs. If it is assumed that trade is not distorted, this situation, hereafter referred to as *Fully decoupled policy* (FD policy), can be used as the benchmark in order to determine if and to what extent a certain type of policy instrument would potentially modify the 'undistorted' production patterns. According to Burell's argument presented in the beginning of this section, there is a degree of ambiguity about the definition of a minimally distorting policy. Hence, it was decided to use the "distance"[1] from the FD policy benchmark as an indicator of distortion potential rather than as an explicit policy objective.

7.2.2.3. Conditional (coupled) Single Payment

The term 'Conditional' here refers to the requirement that a single area payment is provided as long as the land is kept in agricultural use *i.e.*, recipients carry out some activity on the land even though no production of any particular commodity is required. This condition means that the payment is included in the SASEM model's objective function when any of the alternative production activities or set-aside is selected, but it is excluded when land is left idle. Therefore the implementation of the conditional single payment in the SASEM model requires that each hectare that receives a payment must be associated with a current hectare of land. This also resembles the current situation. However, the exact existing situation for the whole agricultural sector is hard to model since the single payments are farm specific and an average national single payment is hard to estimate. It is assumed that this average value is likely to fall within the range between £0/ha and £300/ha for every hectare of land in agricultural use. The APOLO model finds the set of single payment rates (in the decision variable space) that correspond to the Pareto-optimal solutions in the objective space for each scenario examined.

7.2.2.4. Land use subsidies

The proposed choice of policy variables is based on a variant of the arable area payment scheme (AAPS) used in Scotland prior to the CAP reform of July 2003. In order to respect the decoupling requirement of support and at the same time to allow

[1] This distance in the present application really refers to the differences in the levels of the various production activities.

for the promotion of the fittest[2] activities, only area payments were considered. Thus, all cropping activities (cash and fodder crops) are eligible for a per area payment whereas per head payments to all livestock activities are excluded. Specifically, each cropping activity considered in the model can be subsidised at a different rate. This hypothetical scheme is called land-use policy instrument (LUPI) and consists of a number of different per ha payments, one for each of the different land-use activities (cropping activities) represented in the SASEM model. The exact number therefore, of the elements of the LUPI depends mainly on two factors; first, the level of differentiation of the support required by the policy planners and second, the degree of detail with which the possible land-use options can be represented.

7.2.2.5. Taxes on inputs (Nitrogen)

Taxing nitrogen emissions from agriculture is a very difficult task, because emissions are both spatially very diverse and consist of different chemical compounds. Taxing the input of nitrogen (N) in fertilisers is a much simpler policy alternative. Therefore, a percentage price levy per Kg of N Fertiliser is used, which addresses intensity of Nitrogen use on a per-hectare basis. The percentage ratio ranges from 0% to 200% of N fertiliser price per Kg. It should be noted that it is an objective-specific instrument.

7.2.2.6. Policy regulations/constraints

Four regulatory constraints were assumed to be operational and meaningful in the SASEM model: A minimum requirement of set-aside[3] (MSA) area nationally, a maximum quota of finished calves (QFC), a maximum allowable amount of N fertiliser usage (TNU) nationally, and a maximum national level of total soil erosion susceptibility due to land cover (TE_c). Allowable ranges for their values have to be assigned otherwise the model can produce acceptable (feasible mechanically) but not

[2] Again, the term is used here in line with the definition of fittest policy solution given in chapter six.
[3] If the rules for set-aside are relaxed to provide more flexibility *e.g.* in relation to strip widths for conservation, set-aside can be seen as a measure to deliver environmental objectives.

realistic (feasible in practice) solutions. For example, if zero is allowed as a lower value for TNU the model will give the feasible solution where all activities that use N fertiliser (and those depending on them) have zero levels. In practice though, this solution is unlikely to be observed.

Table 7.3. Policy instruments

Type of Instrument	Instrument(s) Description & units	Instrument Code	Range
Unconditional Payments	Unconditional Single Payment (£)	USPI	
Conditional Single Payment	Conditional Single Payment (£/ha)	CSPI	0-300
Land Use Direct Payments	Land Use Subsidies (£/ha)	LUPI	0-300
N Fertiliser Taxation	Percentage Levy on per Kg Price of synthetic N Fertiliser	TI	0-200%
Regulatory constraints	1. Min Requirement of Total Set-aside Area (ha)	RI	$(0\text{-}3)\text{x}10^5$
	2. Number of Finished Calves Quota (hd)		$(3\text{-}5)\text{x}10^5$
	3. Max Allowable Total N Fertiliser Use (TNU) (t/ha)		$(1\text{-}2.5)\text{x}10^5$
	4. Max Allowable Total Erosion Susceptibility due to land cover TEc (t/ha)		$(1\text{-}2.5)\text{x}10^5$

7.3. Farmers' Model

Chapter four is devoted to the standard specification of SASEM, a PMP-based economic model for the agricultural sector of Scotland. Due to the specificities of the PMP method the structure of the standard model (especially in terms of deciding which production activities to include) heavily relies upon the availability of census data. In other words, only activities for which base year observations were recorded in the June censuses or estimated indirectly (see section 4.3) could be calibrated and included in model. The resulting level of detail of the standard model however, does not always suffice to appropriately investigate all scenarios required by a certain policy analysis. This is a considerable drawback of PMP-calibrated models particularly when the policy analysis aims to examine the impact of commercially new activities (not widely adopted by farmers and thus not recorded as a separate category in the agricultural census) on the system under study. Consider, for example, the case where the aim of the policy analysis is to find the mix of activities that overall achieve a reduction of soil erosion potential (as is the case in many of the scenarios presented earlier). If the model's activities are not disaggregated enough in terms of soil erosion potential to represent a range of possible options, there will not be enough (or perhaps any at all) alternative solutions for the model to choose the optimum one, and a feasible set will not be explored properly.

In the case of the analysis conducted in this study, ideally the set of the various production activities should include as many as possible different (in relation to the policy scenarios) farming methods and technologies used by the farmers. In order to expand the standard SASEM model to include these alternative production options that the farmers are faced with, a number of assumptions had to be made and specific techniques used to incorporate these into the SASEM. These are described in the paragraph that follows.

7.3.1. Incorporation of variant activities

Howitt (1995) shows that the nonlinear calibration can take place at any level of aggregation, allowing an LP subcomponent to be nested within the nonlinear objective function of a PMP-calibrated model to obtain the optimum solution (to the full problem). This can be used particularly in technology selection where different specification of, for example, N fertiliser applications or soil protection methods causes discrete production choices.

For the purpose of the present analysis it is assumed that some of the crop activities in the SASEM model can be further differentiated with respect to the amount of N fertiliser applied. For each of these activities two choices are considered: a standard practice and a low N input practice. These alternative choices of the same activity, which hereafter are referred to as variants are listed in table 7.4.

Table 7.4. Variant activities in the SASEM model

Cash Crops	Forage Crops
Spring Wheat (160 Kg/ha N) & (80 Kg/ha N)	Permanent Grassland for grazing (125Kg N/ha) & (175Kg N/ha)
Winter Barley (180 Kg/ha N) & (90 Kg/ha N)	Rotational Grassland for grazing (125Kg N/ha) & (150Kg N/ha)
Spring Barley (100Kg/ha N) & (50 Kg/ha N)	Permanent grassland Hay (125Kg N/ha) & (200Kg N/ha)
Winter Oats (120 Kg/ha N) & (60 Kg/ha N)	Rotational grassland for Hay (125Kg N/ha) & (175Kg N/ha)
Spring Oats (80 Kg/ha N) & (40 Kg/ha N)	Permanent grassland for Silage (125Kg N/ha) & (220Kg N/ha)
Winter OSR (185 Kg/ha N) & (90 Kg/ha N)	Rotational grassland for Silage (125Kg N/ha) & (220Kg N/ha)
Spring OSR (110 Kg/ha N) & (55 Kg/ha N)	

These variant activities compete strongly with each other because they have the same production characteristics apart only from the amount of input N fertilizer. Hence, they have different gross margins as a result of different yields due to different N ferliliser input. However, it should be stressed that the yields were only provisional estimates given that 1) yields depend on a number of (site-specific) factors other that N input for which data or relevant models were not available, and 2) The model assumes same quality and climatic conditions for all the agricultural land.

In the SASEM model they are treated as perfect substitutes. This is illustrated in Table 7.5 which presents a tableau of the links between a certain activity, say activity1, and its two variants Activity1(low) and Activity1(high). Activity1 now has zero gross margins and the same nonlinear variable costs. The gross margins, $Gm_{1(low)}$ and $Gm_{1(high)}$ of the two variants dictate that only the one with the highest gross margins can be selected to enter the optimum solution (one of the variants). The balance row dictates that the level of either variant be less than or equal to the level of activity1. Hence, the latter is forced to enter the optimum solution at the same level which, as chapter 4 shows, overall is determined by its variable costs and the gross margins of the selected variant.

Table 7.5. Incorporation of variants in the SASEM model

	Activity*1*	Variant Activity*1(low)*	Variant Activity*1(high)*	RHS
Gross margin	0	$Gm_{1(low)}$	$Gm_{1(high)}$	
Shadow Variable costs	$-0.5\cdot(q_{11}\cdot X_1^2+\sum^{J_1} q_{1j}X_1X_j)$	0	0	
Balance	-1	1	1	≤ 0

Another assumption is that differences in the N input result in different crop management factor values for grass activities. This assumption was based on the findings of several studies on soil quality that indicate that higher inorganic N in the soil and more frequent mowing both have a negative impact of soil structure potentially resulting in higher risk of soil loss. The estimated *C* values for grassland variant activities typical for Scotland were indirectly derived from Morgan (1995) by consulting soil scientists in Scottish Agricultural College. Table 7.6 presents these values.

Table 7.6. Own estimations of the Mean Annual Values for C for Different

Land Cover	Average Annual Value of *C*
Perm. Grassland for grazing	0.02
Perm. Grassland for Hay (1 cut)	0.06
Perm. Grassland for Hay (2 cuts)	0.07
Perm. Grassland for Silage (2 cuts)	0.08
Perm. Grassland for Silage (3 cuts)	0.09
Temp. grassland for grazing	0.04
Temp. grassland for Hay (1 cut)	0.08
Temp. Grassland for Hay (2 cuts)	0.07
Temp. Grassland for Silage (2 cuts)	0.09
Temp. Grassland for Hay (3 cuts)	0.10
Turnips-Swedes-Rape	0.20
Wheat-Barley-Oats-Oil seed rape	0.25
Set-aside	0.05

SASEM is kept fairly small in size and simple in structure mainly for two reasons. Firstly, to keep the computational burden of the APOLO model within the capability of the programme used, given the fact that for every generation of the MOGA algorithm SASEM is solved as many times as the size of the population of solutions. The second reason is to facilitate the investigation of the relationship between the policy variables and the production variables as well as the identification of desirable policy solutions.

7.4. APOLO Model Formulation

After the specification of the objective functions, the resulting APOLO model can be stated algebraically as following:

Find the vector of subsidies $\mathbf{s} = (S_1, S_2, \ldots, S_j)$

and/or the constraint vector $\mathbf{b} = (B_1, B_2, \ldots, B_j)$

and the activity vector $\mathbf{x} = (X_1, X_2, \ldots, X_j)$

That max $$LanPr = \sum_{j=1}^{J} gm_j X_j / \sum_{j=1}^{J} la_j X_j \quad (7.11a)$$

max $$LabPr = \sum_{j=1}^{J} gm_j X_j / \sum_{j=1}^{J} lb_j X_j \quad (7.11b)$$

min $$MNU = \sum_{j=1}^{J_1} n_j X_j / \sum_{j=1}^{J_1} X_j \quad (7.11c)$$

min $$ME_c = \sum_{j=1}^{J_1} C_j X_j / \sum_{j=1}^{J_1} X_j \quad (7.11d)$$

min $$PolC = \sum_{j=1}^{J} S_j X_j \quad (7.11e)$$

$$PolR = \sum_{j=1}^{J} gm_j (1 + T_j) \cdot X_j \quad (7.11f)$$

Subject to $S_{min} < S_j \leq S_{max}$ *all j = 1 to J*

$T_{min} < T_j \leq T_{max}$ *all j = 1 to J*

max $$TGM^s = \sum_{j=1}^{J} (gm_j (1 + T_j) + S_j) X_j - \sum_{j=1}^{J} \sum_{i=1}^{J} \frac{1}{2} q_{ji} X_j X_i \quad (7.11e)$$

Subject to $$\sum_{j=1}^{J} a_{kj} X_j \leq B_k, \quad \textit{all k = 1 to K} \quad (7.11f)$$

$X_j \geq 0,$ *all j = 1 to J*

This is the standard specification of the APOLO model. However, different policy scenarios are composed of different numbers and combinations of objectives as well as of different types and numbers of policy instruments, requiring the model to be specified accordingly. Therefore, when the model is run under different policy scenarios only the relevant objective functions and variables are included.

7.5. Summary

This chapter demonstrated how the policy scenarios considered for analysis using the APOLO model were designed. First, the political and socio-economic factors shaping the landscape upon which the farming systems in Scotland operate were taken into consideration. Then it was argued that agricultural policy should be designed in a way which ensures that the social goals of multifunctionality and sustainability are achieved in a minimally trade-distorting way. The formulation of policy scenarios involved choosing appropriate policy objectives and instruments. Five policy objectives reflecting relevant indicators of sustainability and multifunctionality were selected for the policy scenarios. It was assumed that three socio-economic and two biophysical indicators represent adequately the socioeconomic and the ecological state of agricultural systems respectively. Also, four sets of policy instruments were selected for the policy scenarios: three sets of economic instruments (single payment, land use subsidies and nitrogen taxation) and one set of regulatory instruments (constraints). With the policy objective functions and policy instruments formulated the rest of the chapter was devoted to the specification and algebraic formulation of the *Follower's* and *Leader's* problems within the APOLO model.

8. Optimisation of Scottish Agricultural Policy: Model Results

8.1. Introduction

As outlined in the previous chapter, a total of five key policy objectives and four types of policy instruments were selected to facilitate policy experiments relevant to Scottish agriculture. Following on from this background information this chapter aims is to demonstrate the model's functionality under real world applications and to illustrate the usefulness of its output in assisting policy decisions concerning agricultural sustainability. In addition, the present chapter serves a second but not less important goal that is, to explore the current situation of Scottish agriculture as a whole as well as possible ways for promoting its multiple functions and enhancing its sustainability.

Each possible policy scenario, which consists of a unique combination of objectives and instruments, forms a unique representation of the system under study suggesting also a unique set of *Pareto*-optimal solutions. This is because for every scenario run the concept of *dominance* which NSGA-II uses to rank different solutions apply to a given representation of the problem. This representation is determined by the structure of the scenario *i.e.*, the number and type of objectives as well as the type of instrument. This implies the unavoidable omission of other relevant dynamics and constraints which could have been detected only by adopting a different scale and a different set of observable indicators, and thus objective functions. Given this diverse set of possibilities it is important to look at several ways of conducting the analysis.

There are at least two possible ways to do this. The first is to examine the effect of increasing/changing the number/combination of objectives on the characteristics (objective values, land use, payment/tax/constraint values) of the solutions comprising the Pareto-optimal set. The second is to examine the relative effectiveness of different types of policy instruments in terms of goal achievement and production patterns. Unfortunately, implementing all possible combinations would involve a large number of scenarios for examination. Thus, only a subset was selected for analysis.

All cases including a single objective were chosen because, as explained in more detail in later sections, their optimisation reveals the model's behaviour and provides important information about the "ideal" state of the modelled system as well as about the potential of the policy instruments. All cases including two objectives were chosen too. Although they are not likely to be very realistic, by conducting pair-wise comparisons of objectives their relationships can be explored straightforwardly. In addition to drawing theoretical conclusions regarding relations between objectives, another important purpose of this application is to examine cases more representative of the real world and if possible to make practical suggestions for agricultural policy in Scotland. However, as pointed out in section 5.4.3.1 and explained in greater detail in section 8.5, due to a number of limitations related to what has been known as the *curse of dimensionality*, multi-objective optimisation was limited to include a maximum number of three objectives.

For each policy scenario, the APOLO model is run to solve the corresponding version of problem (7.11a)-(7.11f). As explained earlier on, due to the simultaneous optimisation of multiple objectives the model finds a population of equally efficient solutions (in a single simulation run a set of *Pareto*-optimal solutions is found). As already pointed out in paragraph 5.2.1, no individual solution dominates the other members of the set with respect to the objective under consideration and a choice among these solutions if desired can be based on higher-level information about the problem at hand. Moreover, being a bi-level programming model, APOLO finds the solution for the leader's decision variables namely, the policy instrument values, as well as the solution for the follower's decision variables; *i.e.* the level of each production activity included in the SASEM model. Therefore, for each policy scenario three sets of solutions are obtained: the policy objectives values, the policy variables and, the farmers' production variables. The first and the third are presented in the main body of this chapter using two types of graphs, scatter plots and bar charts respectively. For the latter a removable legend displaying the colours and patterns for all production activities can be found in the pocket placed in the Appendix (page 199). The output for the policy variables is presented in tables however, owing to their large number and size they are placed in an appendix unless otherwise indicated.

8.2. Single-Objective Optimisation

Optimising one objective provides an initial insight in both the relationship between the various objectives and the relative effectiveness of the policy instruments. Table 8.1 lists the relevant policy scenarios by their code names.

Table 8.1. Single-objective policy scenarios

	Type of Policy Instrument Implemented in the Scenario			
Objectives Included in the Scenario	Conditional Single Payment	Land use Payment	N Fertiliser Taxation	Regulatory Constraints
1. Land Productivity	SO1_CSPI	SO1_ LUPI	SO1_TI	SO1_RI
2. Labour Productivity	SO2_CSPI	SO2_ LUPI	SO2_TI	SO2_RI
3. Cost of Policy	SO3_CSPI	SO3_LUPI	SO3_TI	SO3_RI
4. Mean N Use	BO1_CSPI	BO1_ LUPI	BO1_TI	BO1_RI
5. Mean Erosion Susceptib'y	BO2_CSPI	BO2_ LUPI	BO2_TI	BO2_RI

A way of obtaining useful initial information regarding the nature of association between the objectives in the model is to construct the *pay-off matrix*. This is a square matrix, which is generated by optimising each of the objectives separately over the efficient set and then computing the value of the other objectives. In the literature on multi-objective programming, the vector of the elements on the main diagonal of the pay-off matrix is referred to as the *ideal point* indicating the ideal solution where all the objectives achieve their optimum value. When at least two of the objectives are in conflict the *ideal point* is an infeasible solution. Nevertheless, its elements provide useful information as explained below. Tables 8.1 to 8.4 present the payoff matrices for all the cases of different policy instruments.

Table 8.2. Payoff matrix for the case of conditional single payment

	Objective function value				
Policy scenario	*LanPr* (£/ha)	*LabPr* (£/hr)	*PolC* (M£)	*MNU* (t/ha)	ME_c (units/ha)
SO1_CSPI (Max *LanPr*)	**422.48**	10.11	0	0.168	0.113
SO2_ CSPI (Max *LabPr*)	422.48	**10.11**	0	0.168	0.113
SO3_ CSPI (Min *PolC*)	422.48	10.11	**0**	0.168	0.113
BO1_ CSPI (Min *MNU*)	292.94	8.15	50.81	**0.128**	0.098
BO2_ CSPI (Min ME_c)	260.10	7.48	328.4	0.169	**0.093**

Table 8.3. Payoff matrix for the case of Land Use Payment

	Objective function value				
Policy scenario	*LanPr* (£/ha)	*LabPr* (£/hr)	*PolC* (M£)	*MNU* (t/ha)	ME_c (units/ha)
SO1_ LUPI (Max *LanPr*)	**422.48**	10.11	0	0.168	0.113
SO2_ LUPI (Max *LabPr*)	352.27	**10.56**	304.5	0.164	0.119
SO3_ LUPI (Min *PolC*)	422.48	10.11	**0**	0.168	0.113
BO1_ LUPI (Min *MNU*)	308.50	8.92	338.5	**0.094**	0.116
BO2_ LUPI (Min ME_c)	238.90	6.36	497.0	0.202	**0.064**

Table 8.4. Payoff matrix for the case of N Fertiliser Taxation

	Objective function value				
Policy scenario	*LanPr* (£/ha)	*LabPr* (£/hr)	*PolR* (M£)	*MNU* (t/ha)	ME_c (units/ha)
SO1_CSPI (Max *LanPr*)	**469.83**	10.32	66.3	0.146	0.130
SO2_PI (Max *LabPr*)	469.83	**10.32**	66.3	0.146	0.130
Policy revenue from taxation is not considered as one of the policy objectives but as an indicator					
BO1_PI (Min *MNU*)	469.83	10.32	66.3	**0.146**	0.130
BO2_PI (Min ME_c)	422.48	10.11	0	0.168	**0.113**

Table 8.5. Payoff matrix for the case of Regulatory constraints

	Objective function value				
Policy scenario	*LanPr* (£/ha)	*LabPr* (£/hr)	*PolC* (M£)	*MNU* (t/ha)	ME_c (units/ha)
SO1_CSPI (Max *LanPr*)	**454.62**	10.08	0	0.157	0.126
SO2_PI (Max *LabPr*)	422.48	**10.11**	0	0.168	0.113
SO3 (Min *PolC*)	0	0	**0**	0	0
BO1_PI (Min *MNU*)	297.80	9.52	0	**0.104**	0.103
BO2_PI (Min ME_c)	305.73	9.50	0	0.125	**0.089**

It can easily be seen that corresponding to each ideal value of an objective (diagonal elements) there are four compromise values (the rest of the column's elements). Thus, for a pair of objectives there are two ideal and two compromise values and two 'difference' values the latter being the difference between the ideal and compromise values. The degree of conflict between two objectives can be investigated by considering the two 'difference' values. For example, in Table 8.3, Land Productivity (*LanPr*) conflicts but only weakly with Labour Productivity (*LabPr*). This is because to achieve the maximum value for *LanPr* (£422.48/ha) the maximum of *LabPr* has to be violated by only 0.45 units. This violation represents 11% of the maximum possible violation which takes place when *Mean Soil Erosion* (ME_c) is maximised. Conversely, to maintain its minimum value *LabPr* has to force *LanPr* to drop 70 units below its maximum value or 38% of the maximum possible decrease which is realised when *Mean Soil Erosion* is minimised. It can also be seen that *LanPr* does not conflict at all with Policy Cost (*PolC*) but has a very strong conflict with *MNU* and ME_c since it is when *LanPr* is maximised that the former two also take values close to their maximum possible values while to optimise policy they need to be minimised. By doing all the comparisons the extent of conflict between all the objectives can be revealed. It should be noted that when N taxation is the policy instrument implemented in the model only ME_c appears to be in some conflict with the rest of the objectives.

In terms of the effect of the type of policy instrument to the objective values and their between relationships, two observations can be made. The first is that the various policy instruments differ with respect to their flexibility. This flexibility is reflected in the range of values that each objective can take. LUPI yields the widest ranges for all objective values and thus is the most 'flexible' instrument. N Fertiliser *Taxation Instrument* (TI) on the other hand, is the least flexible of all policy instruments considered because it gives null ranges for three out the four relevant objectives. The second observation is that level of conflict between the objectives changes across policy instrument types which means different *ideal* points for each case. This point should be taken into consideration when comparing the Pareto-optimal solutions generated by each policy instrument.

Although the *payoff matrices* contain very useful information, the true nature and the degree of competition between objectives that is captured by the model can be more properly investigated when each pair of objectives is optimised simultaneously. This is the subject of section 8.3. Another useful characteristic of the pay-off matrix is that the points on the objective space deriving from its entries can be used as an appropriate benchmark comparison against which the characteristics of the *Pareto*-optimal sets can be evaluated.

Another way to represent a comparison between the five payoff solutions is by a radar coordinate system. For five objective functions a circle is divided into five equal arcs. Each radial line connecting the end of an arc with the centre of the circle represents the axis for each objective function. Since the range of values for each objective function differs significantly, objectives have to be normalised. Here, for each line, the circumference marks the 100% of the optimum objective function value (max or min depending on the objective) whereas the centre marks the zero percentage achievement of the objectives. It can be seen that the ideal solution line joins the five edges together forming a regular pentagon. Clearly, the pentagons representing the other solutions have smaller areas indicating that are inferior solutions. It is intuitive that, the closer to the regular pentagon of the ideal solution a shape of a feasible solution is, the better the solution. Although useful for comparing a limited number of solutions, radar graphs is not a very practical option when the number of alternative solutions is large.

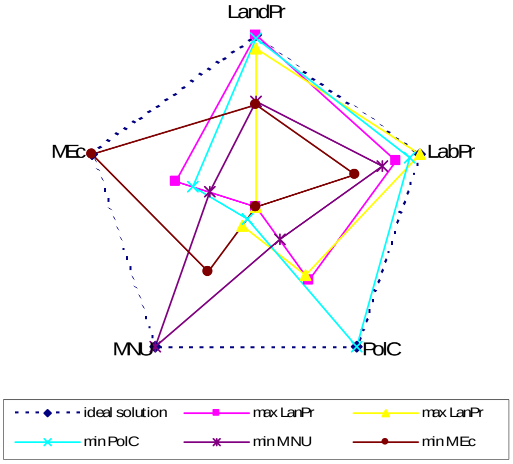

Figure 8.1. A radar graph representation of the payoff matrix solutions for the case of LUPI

The information contained in the objective values is of little practical help to policy makers unless it is supplemented with information about the policy instrument values and the associated production decisions which have produced those objective values. The type and level of production activities in a certain solution partly depends on which policy objective is optimised and which policy instrument is implemented. The relative contribution of a certain activity toward an objective (specified by its policy objective function coefficient) determines its potential to be included in the solution. Its exact level is conditional on the policy instrument in place. For instance, an activity which contributes the least to soil erosion but is also unprofitable can enter the solution if either there is a regulation which poses a minimum requirement for such an activity or there is a subsidy for it which increases its relative competitiveness among the options available to the farmers.

Table 8.6 presents the values of the policy variables for all types of policy instrument that are derived from each single-objective optimisation. The information contained in the table in combination with the information presented in the bar charts of Figure 8.2 can be used to understand the potential of a policy measure to serve a given policy goal. More specifically, it can be examined to determine to what extent a policy instrument can introduce the change in production practices best suited to any given policy objective.

When CSPI is the only measure in effect, optimisation of *LanPr*, *LabPr* or *PolC* results in the same farmers' response *i.e*. land use and thus the same objective values. High N crop activities are carried out because they are more land and labour productive and give the highest return to farmers who are assumed in the model to maximise their gross margins. The optimal level of payment in this case is zero (see Table 8.6). To minimise *MNU*, APOLO determines the single payment rate which stimulates farmers to switch from high N grass activities to their low N variants. At the same time this is accompanied by an increase in permanent grassland area in order to meet the demand for forage. It should be pointed out that the optimum level of the single payment should be such that the mix of land use activities selected corresponds to the minimum possible combined average input of N fertiliser. From Table 8.6 it can be seen that the optimum single payment rate is £36/ha.

Table 8.6. Single objective optimisation results for policy variable values by type of policy instrument and by policy objective

Policy instrument	Objective				
	LanPr	*LabPr*	*PolC*	*MNU*	ME_c
CSPI (x 1)	£0/ha	£0/ha	£0/ha	£36/ha	£180/ha
LUPI (x 30)[1]	All 0£/ha	(**1**:40, **2**:300, **3**:300, **4**:300 **5**:100, **6**:300, **7**:20, **8**:240, **9**:140, **10**:300, **11**:40, **12**:0, **13**:20, **14**:140, **15**:60, **16**:40, **17**:20, **18**:80, **19**:0, **20**:60, **21**:0, **22**:40, **23**:300, **24**:20, **25**:20, **26**:20, **27**:200, **28**:160, **29**:0, **30**:40) £/ha	All 0	(**1**:300, **2**:180, **3**:300, **4**:100, **5**:300, **6**:120, **7**:300, **8**:0, **9**:300, **10**:20, **11**:300, **12**:100, **13**:300, **14**:100, **15**:300, **16**:0, **17**:140, **18**:20, **19**:0, **20**:0, **21**:300, **22**:100, **23**:140, **24**:160, **25**:0, **26**:40, **27**:80, **28**:20, **29**:300, **30**:0) £/ha	(**1**:300, **2**:0, **3**:0, **4**:0, **5**:0, **6**:0, **7**:0, **8**:0, **9**:0, **10**:0, **11**:20, **12**:0, **13**:20, **14**:40, **15**:40, **16**:0, **17**:300, **18**:40, **19**:220, **20**:20, **21**:0, **22**:40, **23**:60, **24**:220, **25**:160, **26**:80, **27**:300, **28**:300, **29**:0, **30**:0) £/ha
RI (x 4)	(0ha, 464420head, 106890 (t/ha), 101470 (t/ha))	(0ha, 388560head, 144720(t/ha), 109820(t/ha))	All 0	(300000ha, 366275head, 236360(t/ha), 100000(t/ha))	(300000ha, 456990head, 100000(t/ha), 137390(t/ha))
TI (x 1)	200%	200%	0%	200%	0%

[1] the 30 crop & grass activities payments refer to: **1**.Set-aside, **2**.Spring Wheat (160 Kg/ha N), **3**.Spring Wheat (80 Kg/ha N), **4**.Winter Barley (180 Kg/ha N), **5**.Winter Barley (90 Kg/ha N), **6**.Spring Barley (100Kg/ha N), **7**.Spring Barley (50 Kg/ha N), **8**.Winter Oats (120 Kg/ha N), **9**.Winter Oats (60 Kg/ha N), **10**.Spring Oats (80 Kg/ha N), **11**.Spring Oats (40 Kg/ha N), **12**.Winter OSR (185 Kg/ha N), **13**.Winter OSR (90 Kg/ha N), **14**.Spring OSR (110 Kg/ha N), **15**.Spring OSR (55 Kg/ha N), **16**.Triticale, **17**.Permanent Grassland for grazing (125Kg N/ha), **18**.Permanent Grassland for grazing (175Kg N/ha), **19**.Rotational Grassland for grazing (125Kg N/ha), **20**.Rotational Grassland for grazing (150Kg N/ha), **21**.Permanent grassland for Hay (125Kg N/ha), **22**.Permanent grassland for Hay (200Kg N/ha), **23**.Rotational grassland for Hay (125Kg N/ha), **24**.Rotational grassland for Hay (175Kg N/ha), **25**.Permanent grassland for Silage (125Kg N/ha), **26**.Permanent grassland for Silage (220Kg N/ha), **27**.Rotational grassland for Silage (125Kg N/ha), **28**.Rotational grassland for Silage (220Kg N/ha), **29**.Turnips and Swedes, **30**.Rape

Higher payments result in land management that overall requires more N fertiliser on average. However, when the payment is set at £180 /ha the solution found by SASEM regarding the activity mixture is associated with the minimum mean soil erosion susceptibility, ME_c. This means that soil protection is a more expensive objective than reduction of nitrogen pollution but as Table 8.1 indicates a marginal reduction in ME_c from 0.098 t/ha (level achieved due to *MNU* optimisation) to 0.093 (level achieved due to ME_c optimisation) requires an almost 6.5 fold increment of policy cost and a significant compromise in N loading (from 0.128 to 0.169 t/ha). Permanent conservation forage for silage is selected instead of rotational. With ME_c the only policy objective the optimal payment rate found by the model promotes grassland extensification whereby making livestock farming more profitable. Thus, the number of animals increases almost 20% and 25% for cattle and sheep respectively.

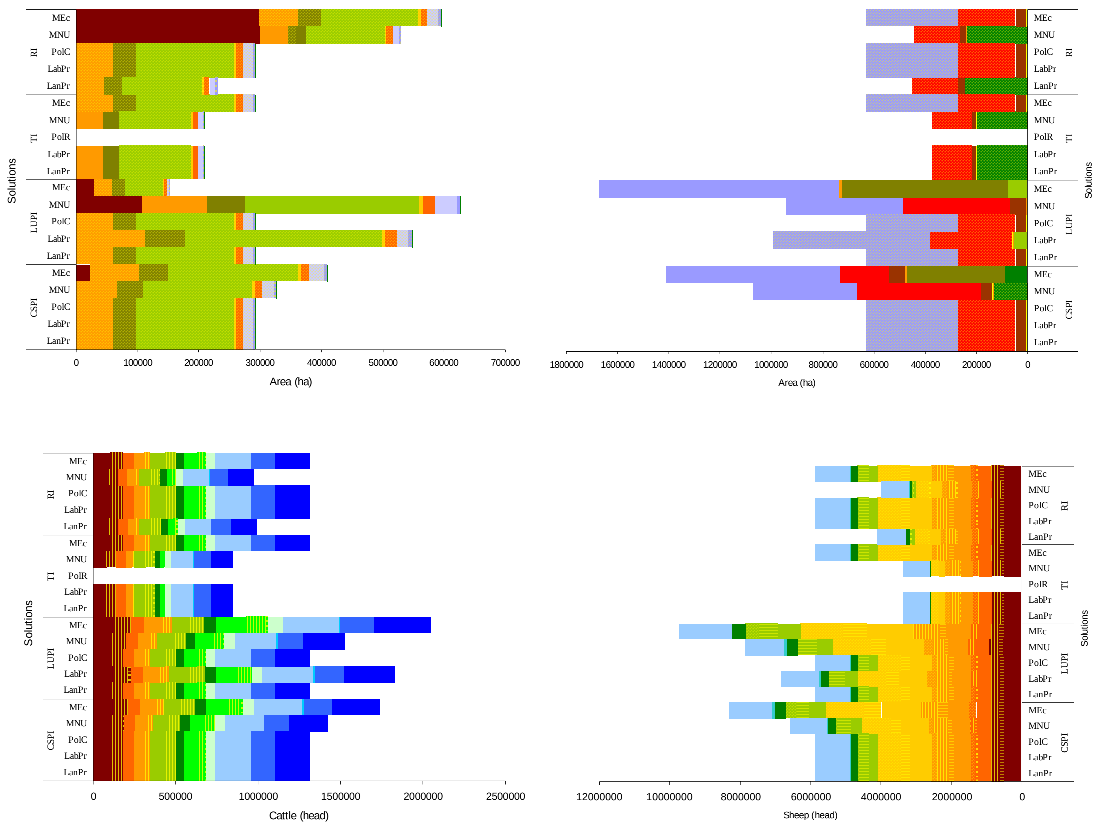

Figure 8.2. Single-objective optimisation results for production activity levels by type of policy instrument

LUPI works in a different way than CSPI. Rather than receiving a single subsidy common for all land activities, under any policy objective a particular activity receives lower or higher subsidy depending on its contribution to the achievement of that objective.

The introduction of any payment has a negative effect on the productivity of land. Hence, if *LanPr* is to be maximised the optimum payment is zero. Consequently, no change in production decisions takes place. As expected, this is also the case when *PolC* is minimised. The situation changes when the policy objective is to minimise *MNU* or ME_c. Minimum *MNU* is realised by switching from high to low N variant activities for all crops. This happens by setting the LUPI subsidies at levels where the low N variants substitute the high N ones in the optimal solution found by SASEM. The set of payments found is such that overall the area of cash crops and set-aside land, which use less N fertiliser than forage crops, increases significantly compared with the baseline scenario. A different set of payments (see Table 8.6) is found which leads to a production pattern associated with the lowest mean soil erosion susceptibility due to land cover, ME_c. In this case land allocated to cash crops and set-aside is almost halved whereas land for grazing and forage production is more than doubled. This is because grassland is less susceptible to soil erosion in the model.

TI is the least flexible policy instrument based on the information contained in the payoff matrix. For N tax rates as high as 200% the high N variants are substituted by the low N variants for only spring wheat and winter barley. Single-objective optimisations with TI give the same output for all cases except for ME_c. In other words, there is no conflict among *LanPr*, *LabPr* and *MNU*. Increasing tax rates and hence the accounting cost of crops can only result in lower profitability of all the cropping activities that have N as an input. Higher taxes force producers to cut production levels, the total land under production is reduced and consequently, more agricultural land is left idle. Also, introduction of N use taxation results in changes in activity levels that are associated with higher soil erosion risk. This is why when ME_c is optimised the optimum tax rate is zero. Although N input taxation can generally result in more efficient fertiliser application, as implemented in the present analysis it is probably not a very effective policy instrument for promoting less intensive use of N fertiliser. For these reasons TI was not considered a well suited instrument for the bi-objective optimisation scenarios.

RI can have a positive effect on all policy objectives except *LabPr*. By keeping set-aside zero and total nitrogen use (TNU) and total erosion susceptibility (TE_c) near their lower limit values *LanPr* slightly improves. This involves reduction in total land under agricultural production. Reduction of *MNU* from 0.157 to 0.104 t/ha (the minimum possible) is facilitated by setting minimum set-aside area (MSA) and TNU constraints to their upper and lower limits respectively.

8.3. Bi-Objective Optimisation

Simultaneous optimisation of two objectives can lead to a set of non-dominated or Pareto-optimal solutions[1]. The position, shape and orientation of the Pareto-optimal sets found by the model depend on the type of bi-objective optimisation. When both objectives are minimised or maximised the Pareto-optimal set extends from a 'min' to 'min' and 'max' to 'max' respectively. When the bi-objective optimisation is of min-max type the Pareto-optimal set extends from a 'min' to 'max' solutions. Table 8.7 shows the most interesting cases of policy optimisation with two objectives covered in this section.

Table 8.7. Bi-objective policy scenarios

	Type of Policy Instrument Implemented in the Scenario		
Objectives Included in the Scenario	Conditional Single Payment	Land use Payment	Regulatory Constraints
1. LanPr – MNU	SO1BO1_CSPI	SO1BO1_LUPI	SO1BO1_RI
2. LanPr - ME_c	SO1BO2_CSPI	SO1BO2_LUPI	SO1BO2_RI
3. LabPr - MNU	SO2BO1_CSPI	-	SO2BO1_RI
4. LabPr - ME_c	-	-	SO2BO2_RI
5. PolC – MNU	-	SO3BO1_LUPI	-
6. PolC – MEc	-	SO3BO2_LUPI	-
7. MNU - ME_c	BO1BO2_CSPI	BO1BO2_LUPI	BO1BO2_RI

8.3.1. Conditional Single Payment Instrument

8.3.1.1. SO1BO1_CSPI (LanPr / MNU with CSPI)

Max *LanPr* and max *MNU* are realised at the base scenario where there is no policy or there is a fully decoupled one. The ideal point is also shown on the scatter plot in Figure 8.3. There is a clear and wide conflict between the two shown by their distant and diametrically opposite optimal points; max *MNU* is realised at max *LanPr*. On the contrary, *LanPr* is not minimum when *MNU* is minimum suggesting that at the point of min *MNU* a value of *LanPr* that is higher than the minimum can be achieved. This is why the Pareto-optimal set originates at the point where *MNU* is minimum. The trade-off curve obtained by the model consists of 26 different non-dominated solutions found out of a population of fifty solutions. It stretches sufficiently from the point of min *MNU* to max *LanPr*. However, solutions are not

[1] In case of two non-conflicting objectives a single optimum solution is found.

uniformly distributed along the curve; most of them are concentrated close to the two extremes.

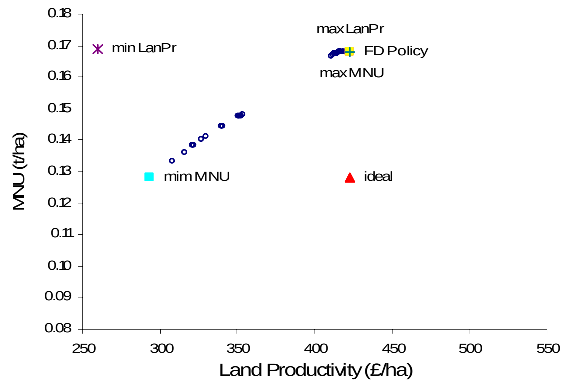

Figure 8.3. Bi-level optimisation results for objective tradeoffs with SO1BO1_CSPI

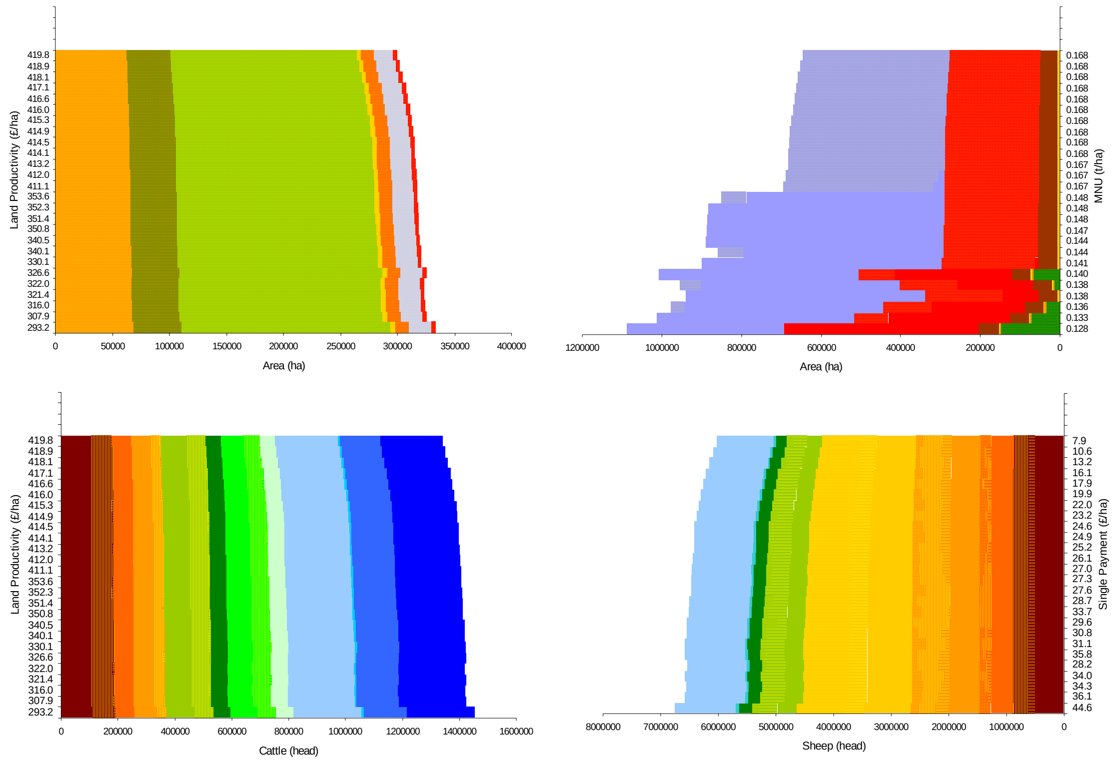

Figure 8.4. Bi-objective optimisation results for production activities with SO1B01_CSPI

The non-dominated solutions found for the two objectives can be obtained only by adjusting the fodder activities; there is no need to change the mixture of the cash crop activities. Moving from the lower left of the non-dominated front to the upper right (Figure 8.3) or equivalently from the bottom to the top bar (Figure 8.4) high N permanent conservation forage for silage production and permanent grassland for grazing activities substitute their low N variant activities. For payments lower than £27 per hectare (14^{th} solution from the bottom) production patterns are almost identical to the baseline situation (zero *PolC* or *FD Policy*).

8.3.1.2. SO1BO2_CSPI (LanPr / ME_c with CSPI)

For this scenario the model was run with a population size of 50 solutions. After 100 generations 6 different non-dominated solutions were found. Of these one is the same as the extreme solution of max ME_c and max *LanPr*. A solution very close the other extreme solution where both objectives take their min values was also found. As Figure 8.5 shows the solutions are not evenly spread across the range of values defined by the two extreme points. This could either be due to the algorithm's failure to find diverse solutions or because all feasible solutions found in the missing range were in fact inferior (dominated) by those finally included in the non-dominated set. Over a number of runs with various different MOGA parameter settings no solutions in the missing range were found indicating that the first explanation is unlikely.

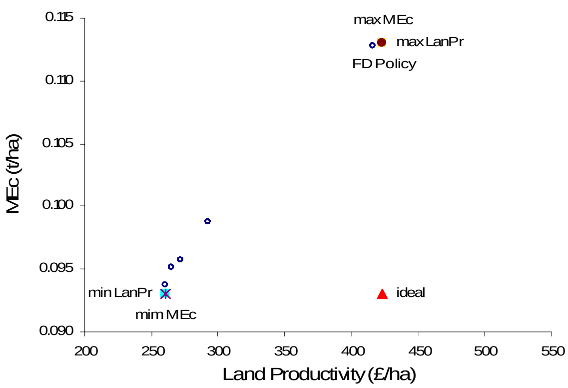

Figure 8.5. Bi-objective optimisation results for objective tradeoffs with SO1BO2_CSPI

The activity combinations which the non-dominated solutions have derived from are shown in Figure 8.6. By selecting activities with lower soil erosion risk (for soil erosion susceptibility factors of the various land uses the reader is referred to Table 7.6) ME_c can be reduced significantly. However, this comes at a higher policy cost, indicated by the higher payments needed since those activities also have lower dry matter yield and consequently are less profitable. To satisfy the demand for livestock forage more area is allocated for forage production compared to the *FD Policy* (baseline) scenario.

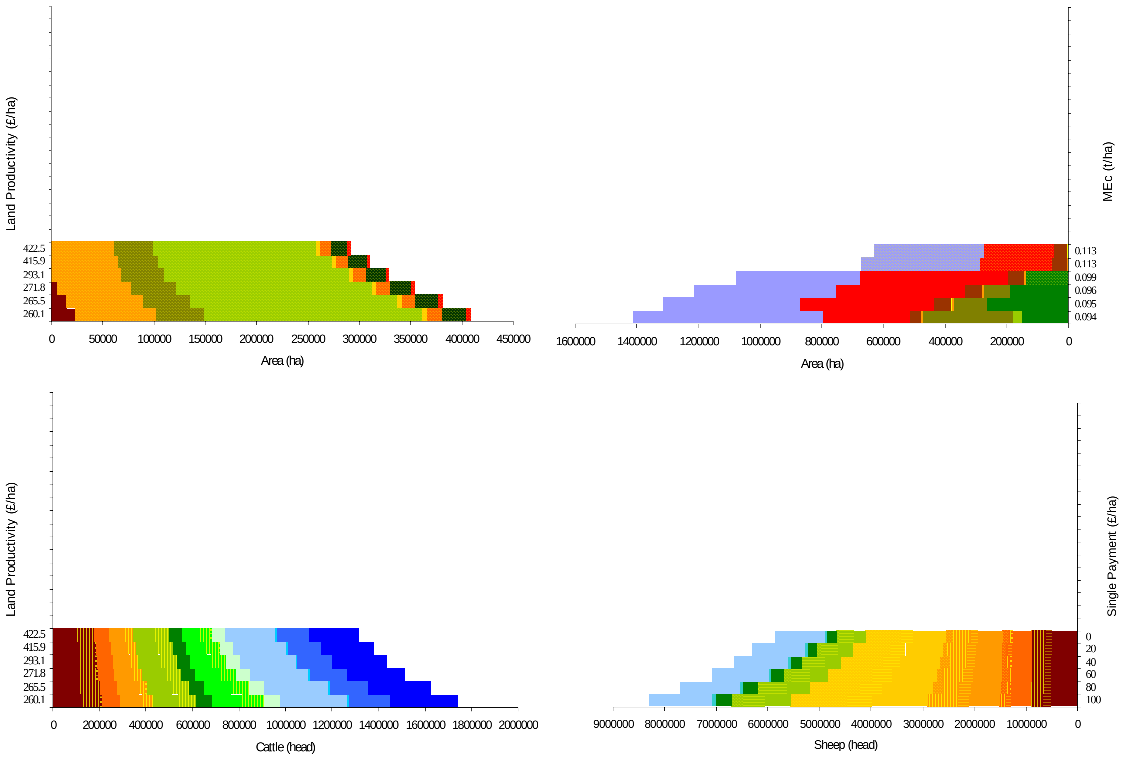

Figure 8.6. Bi-objective optimisation results for production activities with SO1BO2_CSPI

8.3.1.3. SO2BO1_CSPI (LabPr / MNU with CSPI)

There is a negative association between *LabPr* and *MNU* illustrated by the diametrically opposite points of max *LabPr* and min *MNU* (Figure 8.7). The points of max *LabPr* and max *MNU* coincide. However, this is not the case for the other pair (min *LabPr* and min *MNU*); in fact, the latter dominates the former. For population size 50 and 100 generations the model discovered only 3 different non-dominated solutions. Two of these are equal to the extreme points of min *MNU* and max *LabPr* values that are permitted under this specific policy instrument. The reason for this is that all feasible solutions give only these values for *LabPr*. From all these solutions the three found are the non-dominated ones.

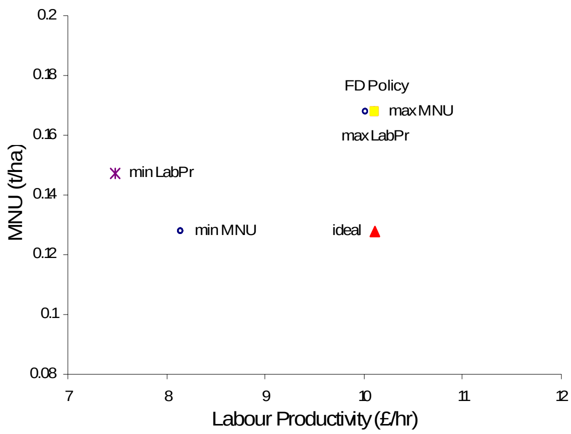

Figure 8.7. Bi-objective optimisation results for objective tradeoffs with SO2BO1_CSPI

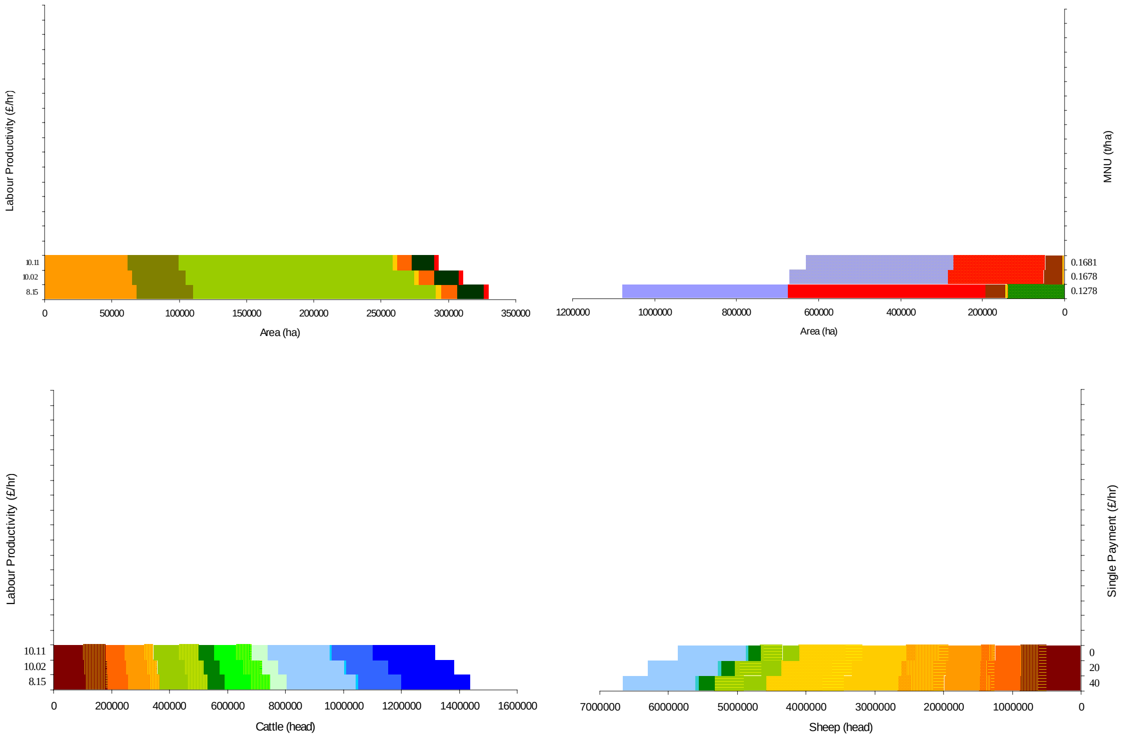

Figure 8.8. Bi-objective optimisation results for production activities with SO2BO1_CSPI

The associated set of activities and their levels are depicted on the bar charts in Figure 8.8. In this case among all arable crops those with low N usage offer the best compromise between the two objectives. For forage crops there are two distinct choices, either extensification (first bar from the bottom) or intensification (the other two bars on top) of grassland and forage production. Both options provide similar supplies of dry matter and hence can sustain roughly the same number of cattle and sheep.

8.3.1.4. BO1BO2_CSPI (MNU / ME_c with CSPI)

Running the APOLO model for this scenario after 100 generations from the 50 non-dominated solutions found only 8 were actually different. These are depicted in the scatter plot in Figure 8.9. The Pareto-optimal front formed by these solutions provides some information about the tradeoffs between the two objectives. More specifically, the opportunity cost of ME_c reduction, measured in terms on *MNU*, can be estimated i.e. how many units of *MNU* have to be foregone for a unit reduction in *MEc*. Moving from the left top to the right bottom, the opportunity cost is decreasing because the slop of the tradeoff curve is decreasing. The bar charts in Figure 8.10 present the mixture of activities and their levels that correspond to the 8 Pareto-optimal solutions.

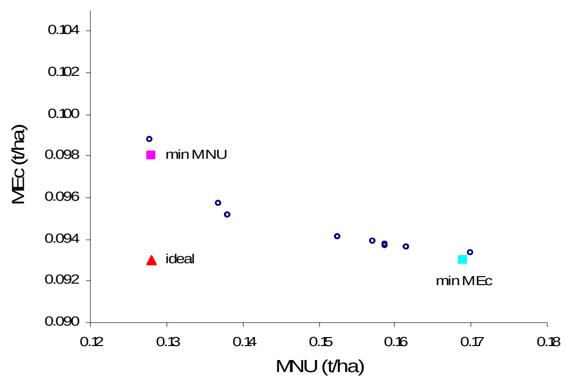

Figure 8.9. Bi-objective optimisation results for objective tradeoffs with BO1BO2_CSPI

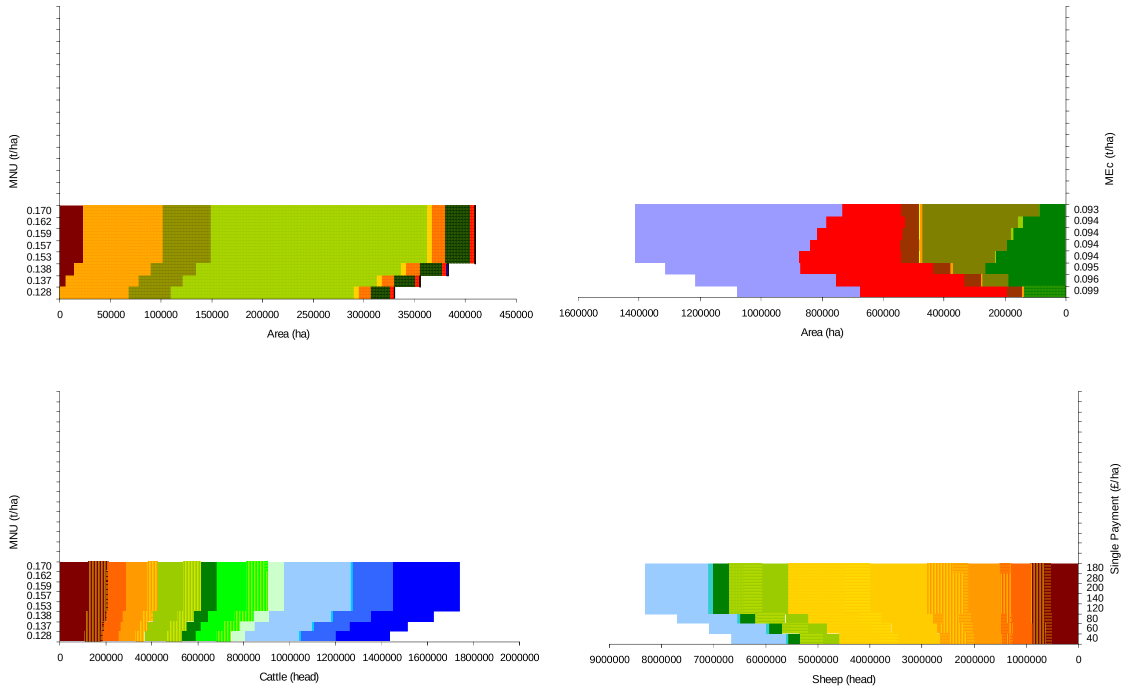

Figure 8.10. Bi-objective optimisation results for production activities with BO1BO2_CSPI

8.3.2. Land Use Payments Instrument (LUPI)

8.3.2.1. SO1BO1_LUPI (LanPr / MNU with LUPI)

The model was run initially with population size 50 solutions for 100 generations. The non-dominated set of solutions did not have a satisfactory spread. Hence, the population size was increased to 200 solutions and the number of generations to 200. Many solutions found were very close to each other in the objective space. Only 26 of them are presented in Figure 8.11. Apart from a small gap near the middle, the set appears continuous with a good spread. Nonetheless, the algorithm failed to find the extreme points or solutions close to them. In fact, it did not converge to the upper extreme Pareto-optimal solution and, solutions found by the model to be members of the non-dominated set are dominated by the min *MNU* extreme solution. The latter case suggests that although the non-dominated set is very close to it is not the same as the true Pareto-optimal set.

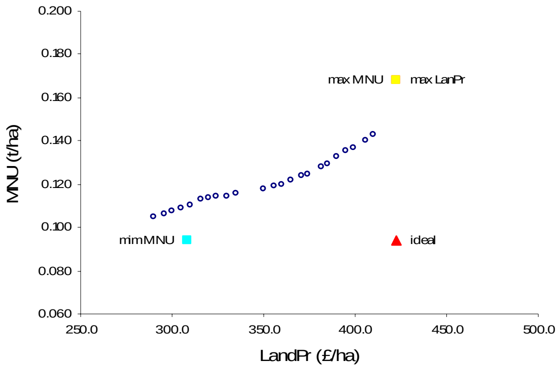

Figure 8.11. Bi-objective optimisation results for objective tradeoffs with SO1BO1_LUPI

This is also reflected in the production activity levels. By referring to the mix of activities associated with the solution where *MNU* takes its minimum value (Figure 8.2), one would expect to see in Figure 8.12 a switch from high N rotational grass for grazing (dotted green bars) to low N rotational grass for grazing. As mentioned above, the algorithm converges to local Pareto-optimal solutions in the region near the extreme points of minimum *LanPr* and *MNU*. As a result rotational grass for grazing 150N is selected instead of permanent grass for grazing 125N. Examining the TGM values reveals that minimum achievement of policy objectives does not coincide with the lowest utility for farmers. In other words, policy makers' and producers' objectives are not completely incompatible.

In terms of cropping and livestock activities a number of changes take place from low to high objective values. These are presented in Figure 8.12. As one moves from low to high objective values *set-aside* area decreases, low N variants of grassland activities are substituted by high N ones and total agricultural land use reduces. These changes represent the response of farmers to the 26 different sets of the LUPI payment values shown in Tables 8.5.1 & 8.5.2 in the appendix. The associated variations in total gross margins levels are also presented in the y-axis of the low bottom bar chart. No clear relationship between TGM and objective values is apparent.

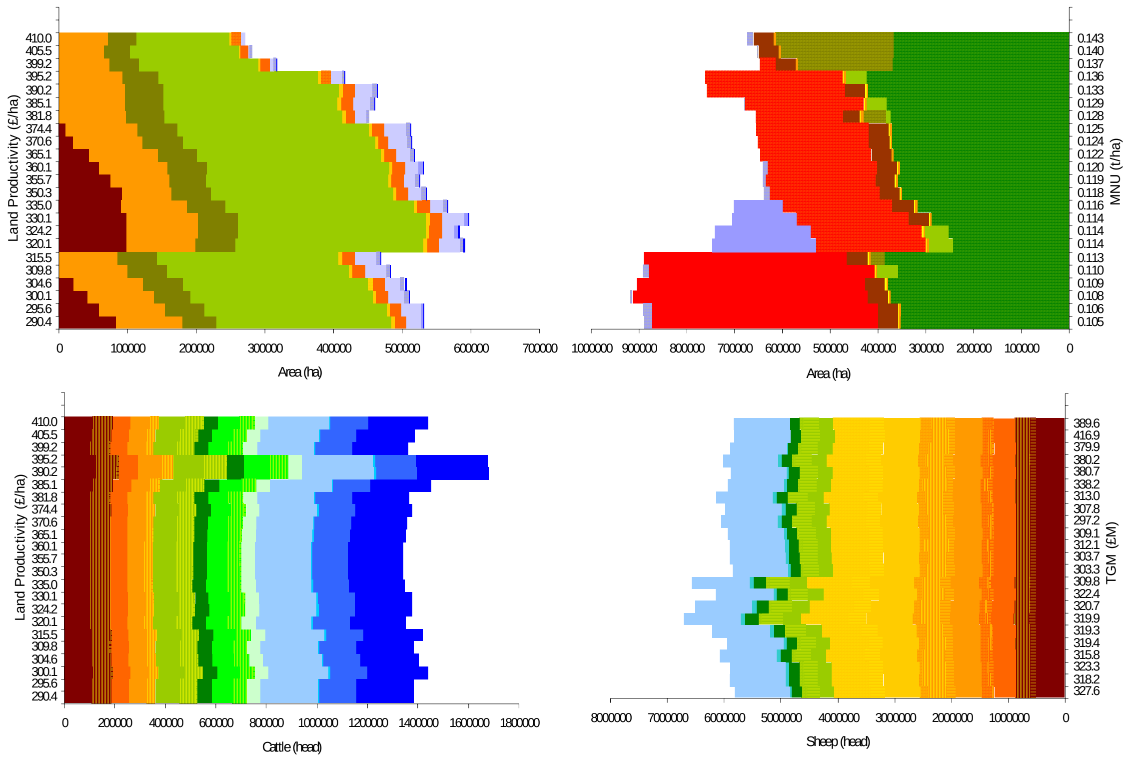

Figure 8.12. Bi-objective optimisation results for production activities with SO1BO1_LUPI

8.3.2.2. SO1BO2_LUPI (LanPr / ME_c with LUPI)

For this policy scenario initially the model was run with a population of 50 individuals for 100 generations. This produced the non-dominated set depicted in the left graph of Figure 8.13. Only 17 representative solutions are shown after removing replicates and solutions that were very similar to those kept. The large distance between the left tail of the set and the lower extreme point suggested that its spread was not satisfactory. After increasing the population to 200 individual and doubling the number of generations the model generated the non-dominated front shown in the right graph of Figure 8.13. As it can be seen the set extends better toward the upper right extreme point and the algorithm this time found a solution superior to the upper extreme point (same max for *LanPr* and lower value for ME_c). Toward the lower left up to a certain point the set has the same stretch as in the first run. However, the algorithm also found a solution which is very close to the lower left extreme point (and non-dominated by it) but no any solutions in between.

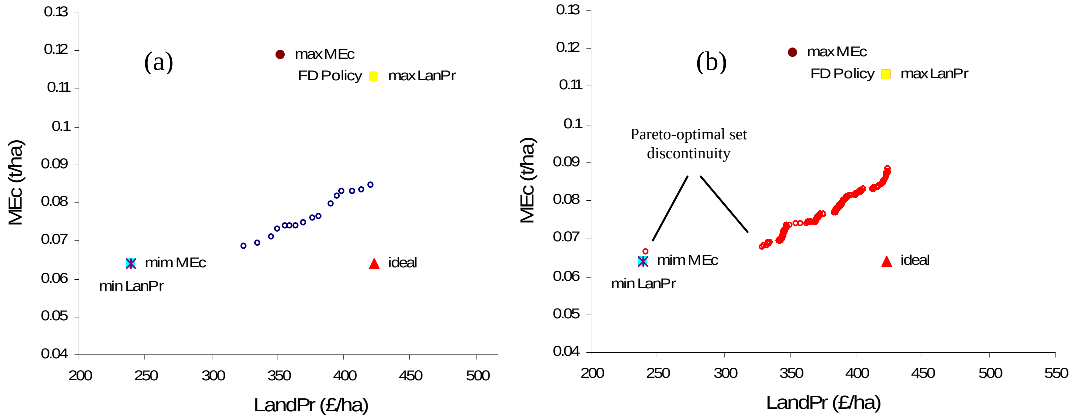

Figure 8.13. Bi-objective optimisation results for objective tradeoffs with SO1BO2_LUPI. Graph (a) shows 17 out of 50 non-dominated solutions found from population size 50 and 100 generations; Graph (b) shows 200 non-dominated solutions found from population size 200 and 200 generations.

This gap in the non-dominated set is an indication that the objective space is discontinuous. As discussed briefly in section 5.4.3.3 discontinuity of the objective space can hinder the algorithm of finding diverse solutions. In such a case, increasing the number of solution in the population and the number of generations can overcome the problem. However, this comes at a higher computational expense.

Figure 8.14 illustrates the set of production decisions associated with the 17 non-dominated solutions. The corresponding payment values which lead to these decisions are given in Tables 8.5.3 & 8.5.4 in the appendix. Compared with the previous scenario changing only one objective, (*i.e.* replacing *MNU* with ME_c) leads to significantly different land use. For example, more than 2/3 of land allocated to cash crops in the baseline scenario is taken out of production. This response by the model is not surprising since these activities can contribute up to 10 times more to soil erosion compared with forage activities.

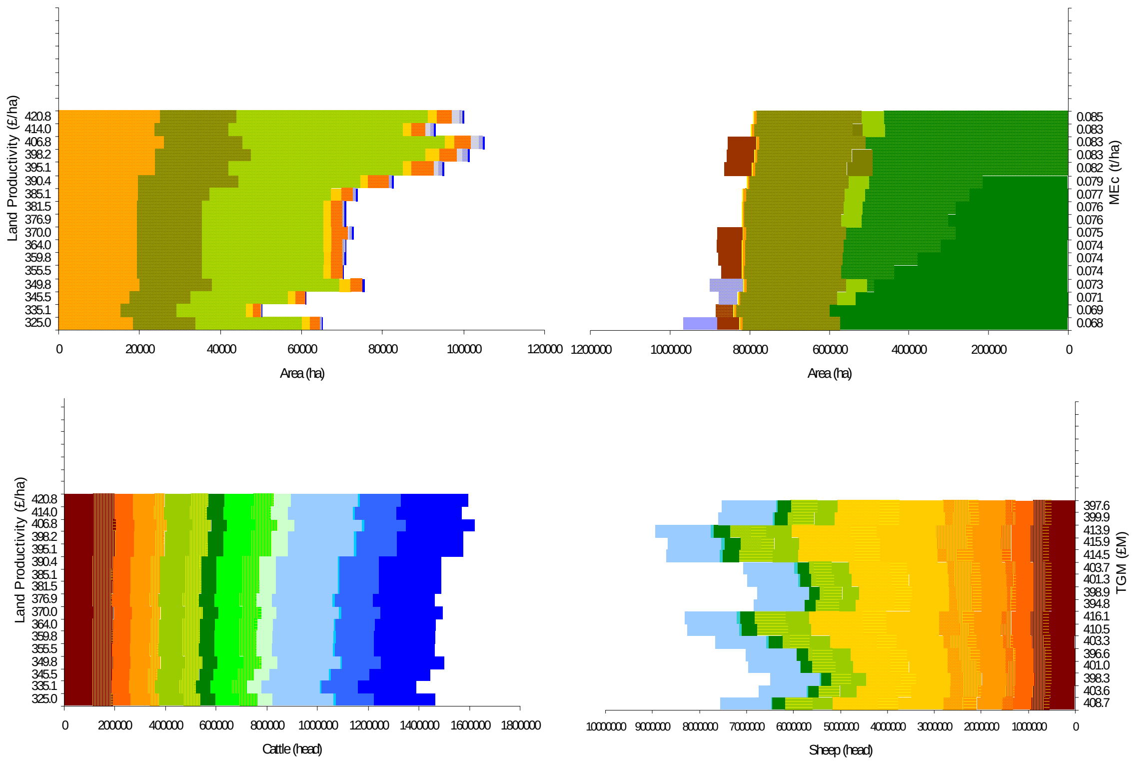

Figure 8.14. Bi-objective optimisation results for production activities with SO1BO2_LUPI

8.3.2.3. SO3BO1_LUPI (PolC / MNU with LUPI)

Running APOLO with a population of 50 solutions and for 100 generations derived the set of non-dominated solutions depicted in Figure 8.15. There was no convergence to the extreme point where *MNU* take its minimum value. Nonetheless, the set was thought to be good enough because it includes a solution which is very close to the minimum *MNU* value. In practice, trying to discover solutions further to the right of the set makes little sense since, the opportunity cost of further minimising *MNU* measured in monetary values (*PolC* units) gets extremely high. Overall, the marginal opportunity cost of *MNU* reduction (the slope of the trade-off curve with respect to the y-axis) increases with *PolC*.

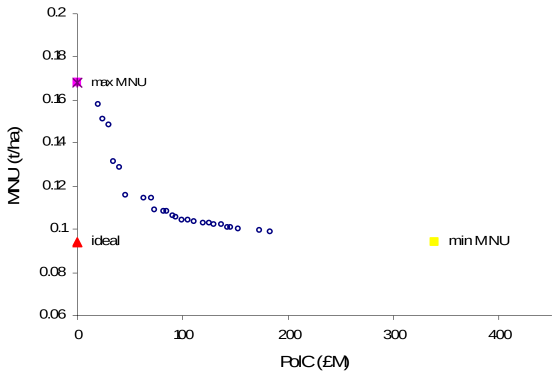

Figure 8.15. Bi-objective optimisation results for objective tradeoffs with SO3BO1_LUPI

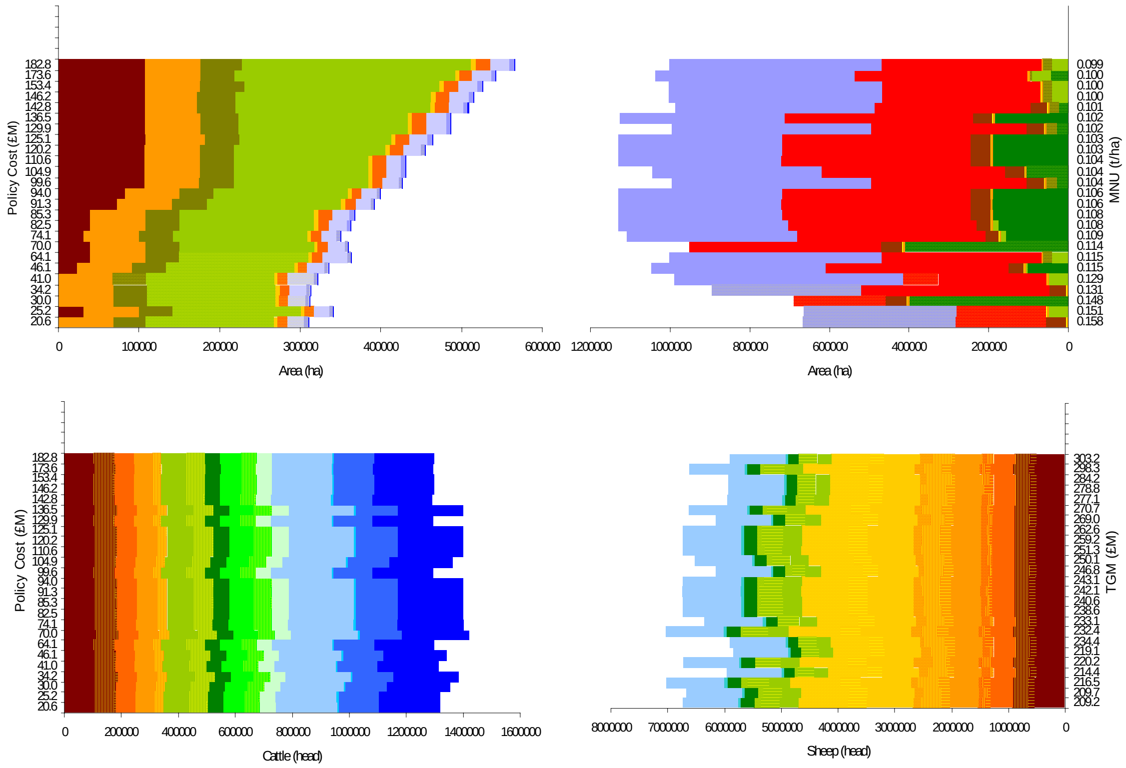

Figure 8.16. Bi-objective optimisation results for production activities with SO3BO1_LUPI

Regarding the land management that generates the above objective values, one interesting, albeit expected, pattern emerges. To decrease *MNU*, low N variant activities have to be selected and for that to happen they have to receive higher payments. Hence as *PolC* increases high N activities are replaced by their low N substitutes. To further lower *MNU*, land-use activities with better overall performance with respect to both objectives, *i.e.* activities that can reduce mean N fertiliser use at the least cost (payment), such as set-aside and spring barley, receive higher payments (see Tables 8.5.5 & 8.5.6 in appendix) and increase their shares.

8.3.2.4. SO3BO2_LUPI (PolC / ME_c with LUPI)

For this scenario Figure 8.17 illustrates the set of 30 out of 50 non-dominated solutions derived from running APOLO with a population of 50 solutions and for 100 generations. In this case too, the algorithm did not converge to the extreme points but, the non-dominated set was thought to be a good approximation of the true Pareto-optimal front. Like in the previous scenario, the marginal opportunity cost (of ME_c reduction this time) increases with *PolC*. In other words, it becomes increasingly more costly to reduce soil erosion by an additional unit.

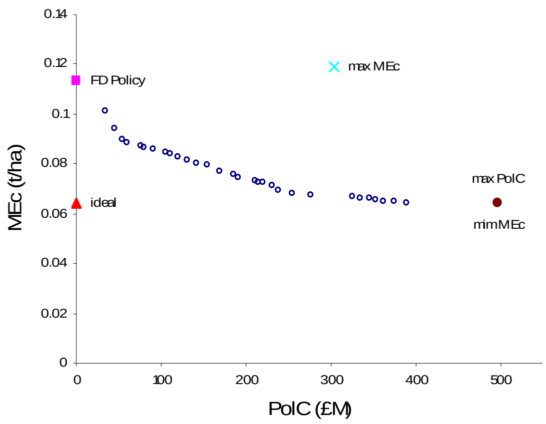

Figure 8.17. Bi-objective optimisation results for objective tradeoffs with SO3BO2_LUPI

Figure 8.18 shows the 30 Pareto-optimal combinations of various land use and livestock activities. There is no incentive to promote low N variant activities here; thus, only the more profitable cash crops are selected. The presence of set-aside land in the solutions is due to its much lower soil erosion factor value. For the same reason permanent and rotational grassland and forage production activities with lower (two) number of cuts receive higher payments in order to substitute the more productive (three cuts) but also more susceptible to soil erosion grassland and forage activities. Further payment increments to permanent grassland result in further reduction of mean soil erosion and at the same time make livestock farming more profitable. Consequently the cattle and sheep sectors expand significantly. However, over these steps policy cost increases exponentially.

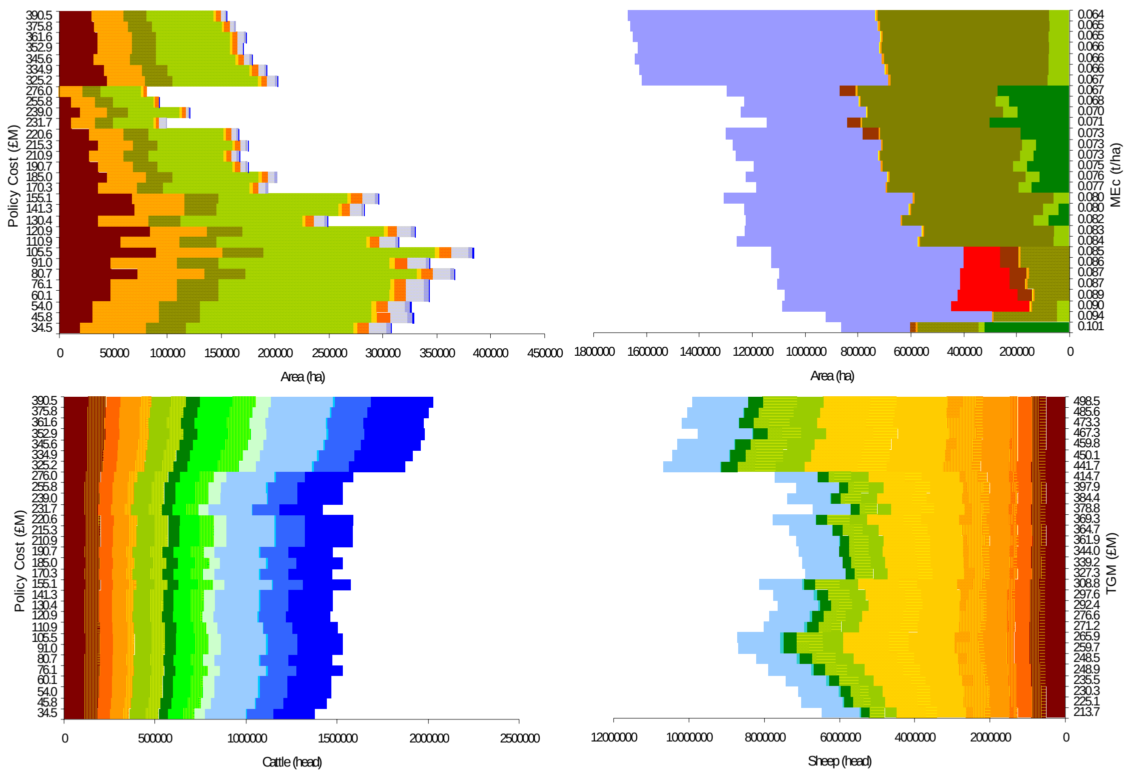

Figure 8.18. Bi-objective optimisation results for production activities with SO3BO2_LUPI

8.3.2.5. BO1BO2_LUPI (MNU / ME_c with LUPI)

For this scenario APOLO was run with the same parameter settings for NSGA-II *i.e.* 50 solutions and 100 generations. The non-dominated set found contained a number of similar solutions. For ease of presentation only 30 representative solutions are shown in the scatter plot in Figure 8.19 below. It can be seen that there are two distinct trade-off regions. The region on the left suggests that the most significant reduction in ME_c can be achieved with a relatively limited compromise in terms of *MNU*. Further reductions are only marginal and require disproportionately higher levels of MNU sacrifice as shown by the smaller trade-off curve region on the right hand side of the graph.

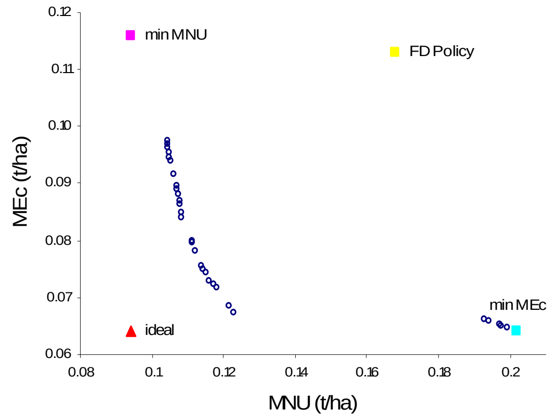

Figure 8.19. Bi-objective optimisation results for objective tradeoffs with BO1BO2_LUPI

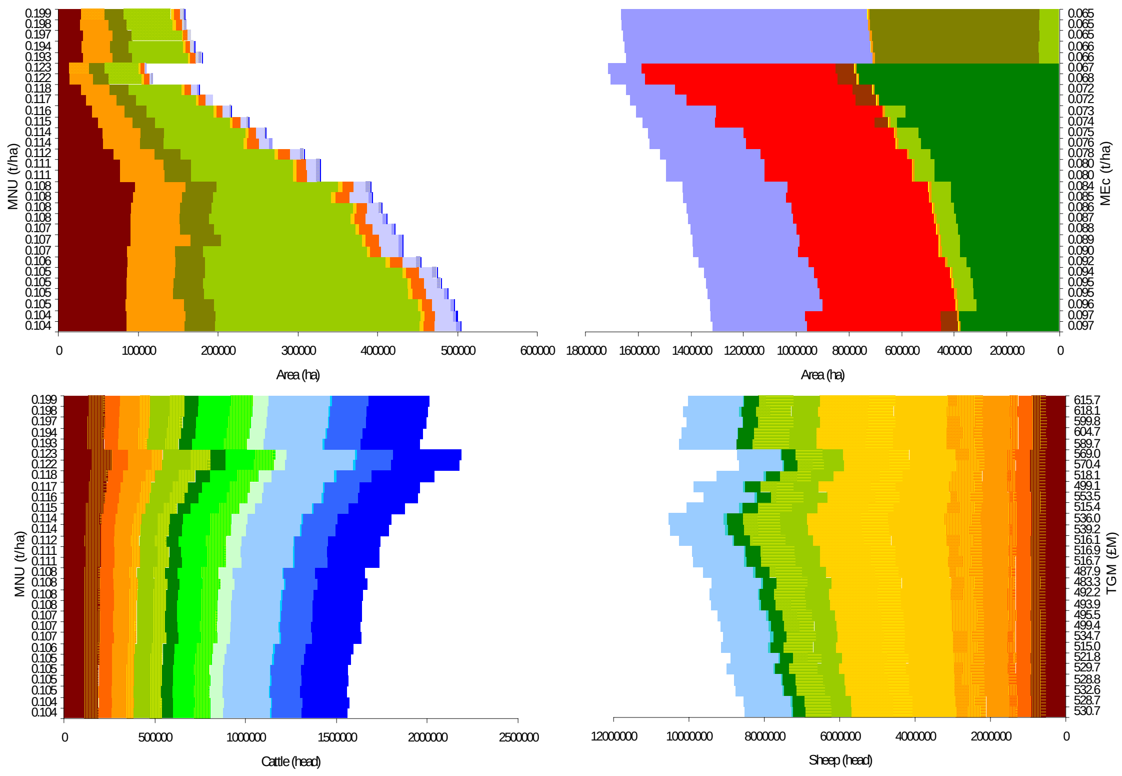

Figure 8.20. Bi-objective optimisation results for production activities with BO1BO2_LUPI

For any given solution in the non-dominated set there is a set of payment values responsible. These are presented in Tables 8.5.7 & 8.5.8 in the appendix. The associated responses anticipated by farmers according to the SASEM model are reflected in the production activity mixtures depicted in Figure 8.20. The main characteristic of this set of land use solutions is the extensification of production both in terms of N fertiliser application and in terms of total area under agricultural use.

8.3.3. Regulatory Constraints

8.3.3.1. SO1BO1_RI (LanPr / MNU with RI)

After 100 generations the model found 30 different non-dominated solutions forming a straight line with a satisfactory spread and well distributed solutions. There is a profound conflict between the two objectives which remains constant over the range of objective values. The corresponding policy variables are presented in Table 8.5.9 in the appendix.

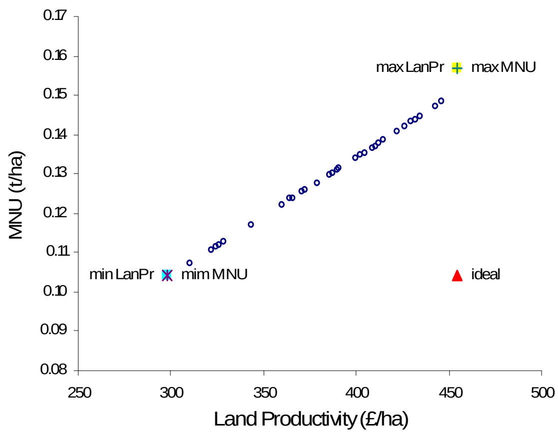

Figure 8.21. Bi-objective optimisation results for objective tradeoffs with SO1BO1_RI

The MNU minimisation objective forces the TNU regulatory constraint to take its lowest allowable (as specified by the characteristics of the RI instrument) value namely, 10^5 tonnes of soil. Consequently, in relation to the baseline scenario, adjustments in the mix of production activities take place. Spring wheat 80N, winter barley 90N and winter oil seed rape 90N substitute spring wheat 160N, winter barley 180N and winter oil seed rape 185N respectively. Also, permanent grass for grazing 175N is replaced by rotational grass for grazing 150N and set-aside land enters the Pareto-optimal land use solutions. The min *MNU* / min *LanPr* values are observed when set-aside area reaches its upper limit. Every hectare of set-aside that is taken out results in an increment of both *MNU* and *LanPr*. This happens because set-aside land receives zero N fertiliser input and has negative *LanPr*. Figure 8.22 illustrates the various production combinations. The variation in TGM is primarily due to the increasing area of set-aside land; the levels of all other activities do not change. While the conflict between the two objectives is evident, combined they force the area of less productive grassland to decrease leading to a significant reduction in livestock numbers both cattle and sheep compared to the baseline levels.

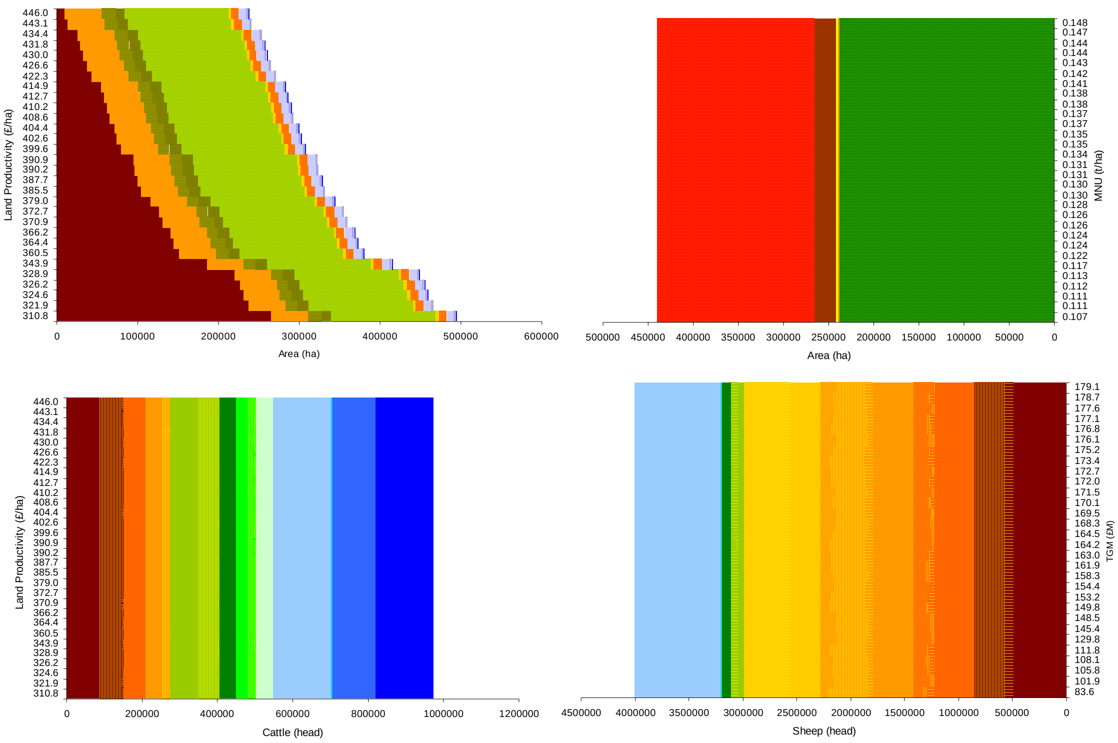

Figure 8.22. Bi-objective optimisation results for production activities with SO1BO1_RI

8.3.3.2. SO1BO2_RI (LanPr / ME_c with RI)

Figure 8.23 shows the trade-off curve which is formed by the 30 different non-dominated solutions found after 100 generations. The conflict between the two objectives remains constant over the range of objective values. The corresponding values of regulatory constraints are presented in Table 8.5.10 in the appendix.

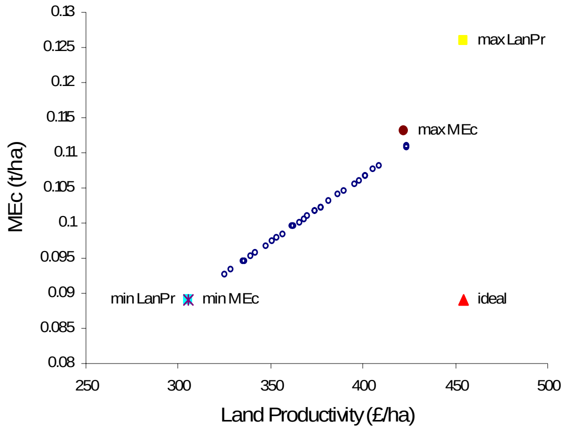

Figure 8.23. Bi-objective optimisation results for objective tradeoffs with SO1BO2_RI

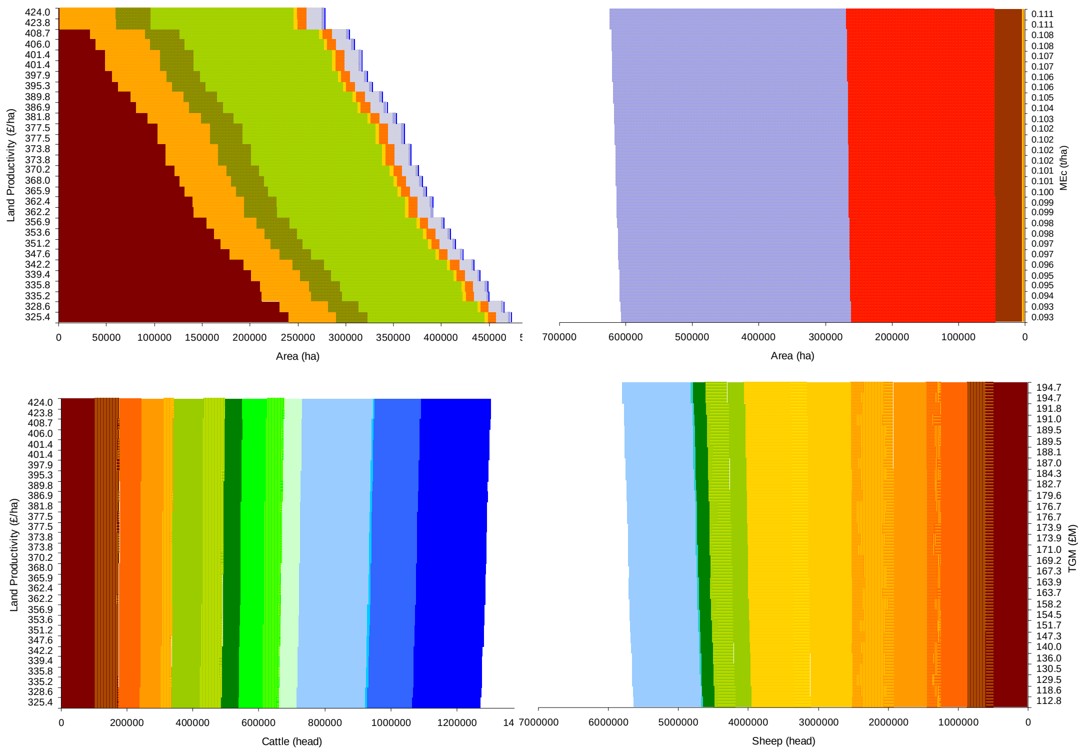

Figure 8.24. Bi-objective optimisation results for production activities with RI (*LanPr - ME$_c$*)

The associated response of farmers to these constraints in terms of production practices and choices as modelled by the SASEM model are depicted in Figure 8.24. Here too set-aside plays a significant role in achieving a range of values for the *ME$_c$* objective. Regarding crop and livestock activities, their shares and their levels remain the same as in the base scenario. Total gross margins fall as the minimum requirement of set-aside area imposed by the policy due to the costs of keeping set-aside land in good agricultural condition.

8.3.3.3. SO2BO1_RI (LabPr / MNU with RI)

Under this scenario the model after 100 generations and population size 30 found equal number of non-dominated solutions shown in Figure 8.25. The trade-off curve which is formed by these solutions suggests that for the largest portion of the objective values the conflict between the two objectives is constant.

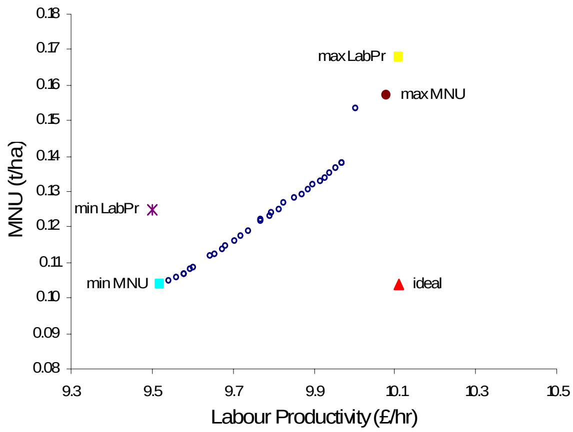

Figure 8.25. Bi-objective optimisation results for objective tradeoffs with SO2BO1_RI

As illustrated in the bar charts in Figure 8.26 this policy scenario results into two distinct production patterns the one associated with the upper and the other with the lower half part of objective values. They differ in that solutions following the second pattern include only spring wheat 80N, both winter barley variants, both winter oil seed rape variants and only rotational grass for grazing 150N. Solutions following the first pattern include only the high N variants of cash crops and only permanent grass for grazing 175N. TGM take their highest and lowest values at max and min objective values respectively. Overall TGM depend on MSA constraint levels. Table 8.5.11 presents the levels of all constraints with the associated objective values for all non-dominated solutions.

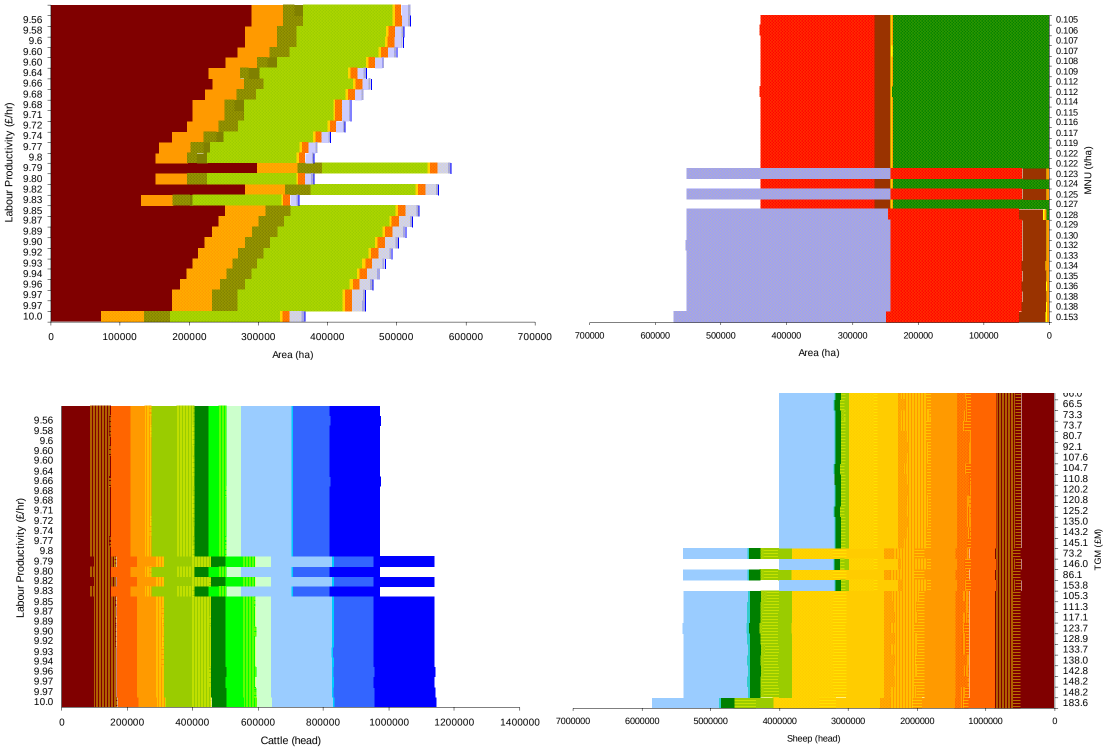

Figure 8.26. Bi-objective optimisation results for production activities with SO2BO1_RI

8.3.3.4. SO2BO2_RI (LabPr / ME_c with RI)

Figure 8.27 depicts the 30 non-dominated solutions found by the APOLO model run for 100 generations and a population size 30. The solutions are well distributed but there is no convergence to the upper extreme solution where both objectives take their maximum values. The slope of the trade-off curve that is formed indicates that the kind of conflict is similar to the previous case.

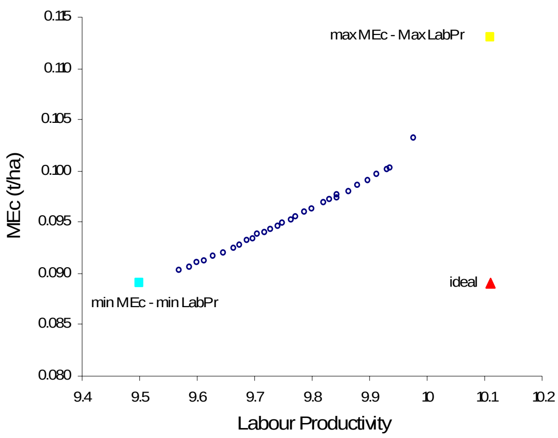

Figure 8.27. Bi-objective optimisation results for objective tradeoffs with SO2BO2_RI

8.3.3.5. BO1BO2_RI (MNU / ME_c with RI)

Running the APOLO model for this scenario with a population of 30 solutions for 100 generations generated the set of 30 non-dominated solutions shown in Figure 8.28. The non-dominated front exhibits a satisfactory diversity of uniformly distributed solutions.

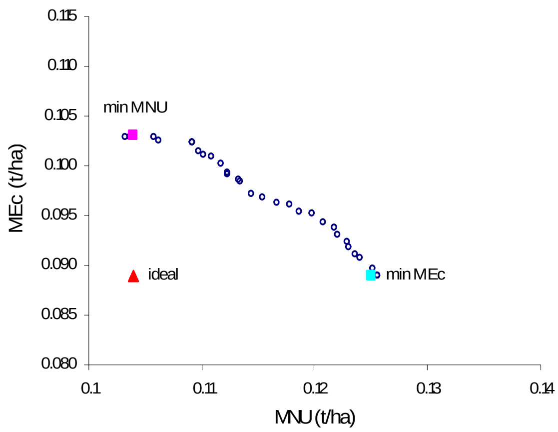

Figure 8.28. Bi-objective optimisation results for *MNU - ME_c* tradeoffs with RI

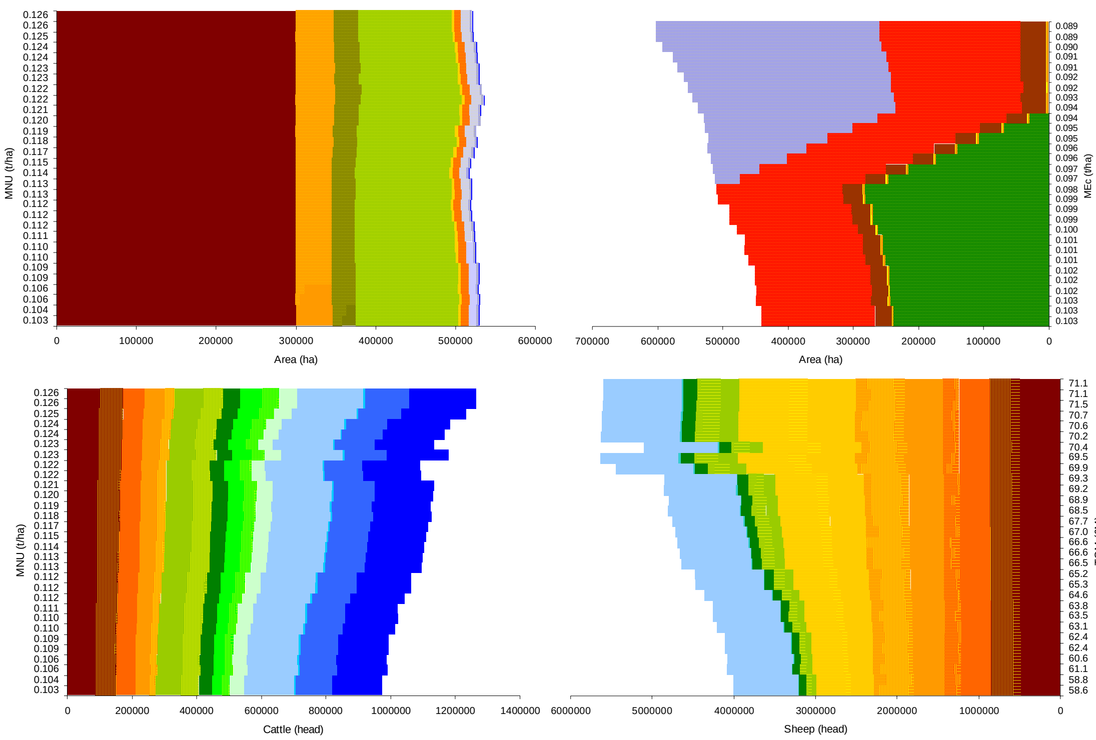

Figure 8.29. Bi-objective optimisation results for production activities with RI (*MNU - ME_c*)

The corresponding levels of MSA, QFC, TE_c and MNU regulatory constraints are given in Table 8.5.13 in the appendix. All solutions use the maximum requirement of set-aside area, which is also reflected in Figure 8.29. As expected, TE_c and TNU levels decrease and increase respectively from solutions of low *MNU* to low ME_c objective values. In terms of production decisions this corresponds to a switch from rotational grass for grazing 150N to permanent grass for grazing 175N, an increase in total grassland area followed by an increase in number of cattle and sheep as well as a rise in total gross margins.

8.3.4. Comparing policy instruments

A comparison of the policy instruments examined in the present application can be in a number of different levels. For example, the concept of dominance could be used to assess which policy instrument generates the most superior solutions. Also, the diversity and number of alternative options are important attributes which could be used to evaluate an instrument's flexibility and efficiency. Other ways based on auxiliary information such as farmers' utility and impact of the production patterns may important decision factors too.

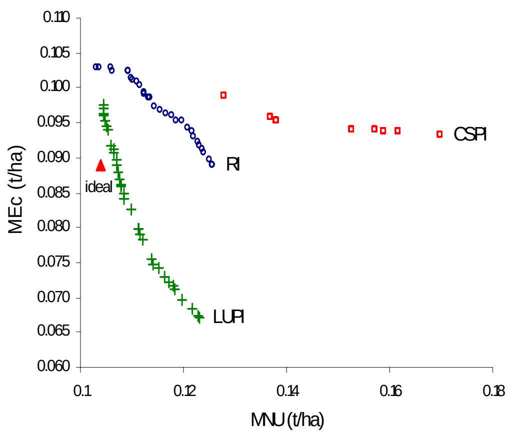

Figure 8.30. Comparison of CSPI, LUPI and RI with respect to environmental objectives

Take the case where policy optimisation concerns only two environmental objectives, *MNU* and ME_c. If other things assumed equal or less important and the comparison is based purely on the objectives attainment then, as shown in Figure 8.30, LUPI finds a set of policy solutions that clearly dominate solutions derived from the other two policy instruments. It can be assuming that for Scotland the current situation (as modelled) with the CSPI being implemented corresponds to one of the CSPI derived solutions. It is obvious that whatever the current situation it cannot be better than the CSPI-derived non-dominated solutions depicted with red squares in Figure 8.30. Therefore, it can be suggested that the single payment scheme currently implemented in Scotland is not the best of all possible policy measures for achieving reduction of N fertiliser use and soil erosion susceptibility.

8.4. Three-Objective Optimisation

8.4.1. Issues for consideration

The previous sections have demonstrated the capacity for the APOLO model to generate sets of non-dominated solutions for a series of representations of the bi-level problem of single and bi-objective policy optimisation for agricultural systems in Scotland. As pointed out already in section 5.4.3.1 the difficulty of finding the true Pareto-optimal front depends on the number of objectives being optimised. When the number of objectives increases, the dimensionality of the objective space also increases. With M objectives, the Pareto-optimal front can be at most an M-dimensional surface. When dealing with a large number of objectives and consequently a high-dimensional search space, one is faced with the ‘curse of dimensionality’ and the associated sparsity issues. Essentially the amount of solutions in the population necessary to sustain a given spatial density increases exponentially with the dimensionality of the search space.

Deb (2001) also shows that an additional implication of increasing the number of objectives is that the proportion of non-dominated solutions in the initial random population increases too. When the initial population is randomly created, this may cause difficulties for most multi-objective evolutionary algorithms because they emphasise all solutions of the first non-dominated front equally by assigning the same or a similar fitness. When this happens, there is no selection advantage to any of these solutions. Consequently, in the absence of any selection pressure for better solutions, the task of crossover and mutation operators to find (create) better solutions may be difficult in general. Elitism will not overcome the problem either simply because most of the population members belong to the best non-dominated front and there may not be any population slot left to include any new solution.

One way to mitigate the problems of dimensionality is for the algorithm to use some other criterion instead of assigning fitness based on the non-dominated rank of a solution. The amount of spread in objective space may be used to assign fitness. In fact the NSGA-II, although it initially assigns exactly the same fitness to all solutions of the first non-dominated front, it then applies the *crowded tournament selection* operator to favour solutions that lie in less dense regions on the non-dominated front compared with those that are closely packed in one part of the non-dominated front. While this operator ensures the selection of more uniformly distributed solutions, it does not provide a mechanism to discover solutions closer to the true Pareto-optimal front.

Another way to alleviate the problems of dimensionality is to increase the spatial density of the search by using a large population size. Deb (2001) has simulated the proportions of best non-dominated solutions for different number of objectives and provides a chart for finding the minimum adequate population size. From that chart it is clear that the required population size increases exponentially with the number of objectives (Deb 2001 pp 419). Alternatively, if information about good regions of the decision space is available, the algorithm's population can be initialised there instead of randomly in the entire decision space. Such information can be provided from Pareto-optimal solutions that have derived from optimisation with a subset of the objectives.

Figure 8.30 shows the Pareto-optimal surface in a typical three-dimensional minimisation problem. All solutions on such a surface are non-dominated to each other. Thus, the feasible search space lies above this surface. The task of the MOGA is to reach this surface from the interior of the search space and distribute solutions as uniformly as possible over the surface. Unfortunately, in real-world applications the true three dimensional Pareto-optimal set may not form a regular and continuous surface as in the above example thereby making the non-dominated solutions appear as randomly distributed. Therefore, pair-wise plots of any two objectives may not result in a recognisable trade-off front.

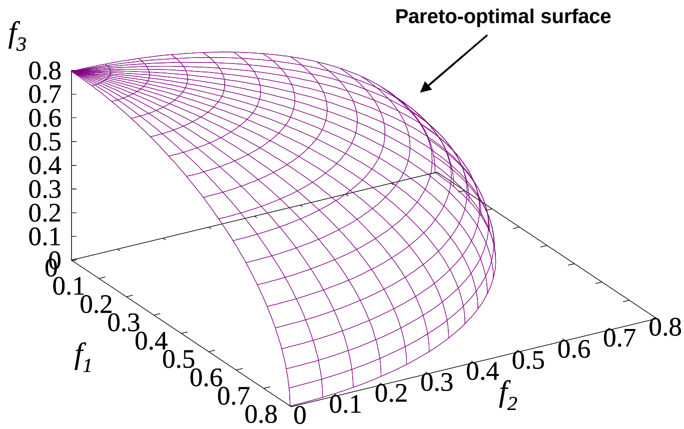

Figure 8.31. A regular Pareto-optimal surface for a three-objective minimisation problem

8.4.2. Policy Objectives Results

The remainder sections of this chapter present the APOLO model output for the policy objective variables obtained from the optimisation of three objectives when the LUPI policy instrument option is in effect. Only four objectives were considered namely, *LanPr*, *PolC*, *MNU* and ME_c. Their combinations gave the three following policy scenarios: SO1SO3BO1_ LUPI, SO1SO3BO2_ LUPI and SO3BO1BO2_ LUPI. The setting of the NSGA-II algorithm's parameters was done empirically taking into consideration the following criteria. (1) the spread and distribution of the non-dominated front obtained at the end of the models' cycle along each objective axis, (2) the proximity of these fronts to their counterparts obtained from the bi-objective optimisations and, (3) the CPU time. After several trials with various parameter settings the NSGA-II population and generation number were set to 500 and 200 respectively for all scenarios; a relatively high mutation probability was also used.

The resulting non-dominated policy solution sets are illustrated in the same way for all scenarios using three-dimensional scatter plots and their three two-dimensional projections. In order to aid with the visualisation and the qualitative assessment of the derived non-dominated fronts, two 3D scatter plots are presented for each case (Figures 8.31, 8.32 & 8.35). Plots (a) show the position of the non-dominated front in relation to the random solutions selected at the initialisation stage of the optimisation. Plots (b) show the position of the non-dominated front derived from the three-objective optimisation in relation to the non-dominated fronts obtained previously from the relevant two bi-objective optimisations. The different solution sets are distinguished by colour.

For the aforementioned scenarios the APOLO model found the sets of 500 non-dominated solutions shown in red in all Figures. Two immediate conclusions can be drawn by inspecting the graphs. The first is that, as shown in plots (a) in Figures 8.31, 8.33 and 8.35, the model improves the initial random guess significantly in all cases. The second is that the border of all non-dominated surfaces derived from the three-objective optimisations (*i.e.* SO1SO3BO1_LUPI, SO1SO3BO2_LUPI, SO3BO1BO2_LUPI) touch the non-dominated fronts derived from the relevant two-objective optimisations (*i.e.* SO1BO1_LUPI and SO3BO1_LUPI for SO1SO3BO1_LUPI etc)[2]. This means that, with the chosen parameter specification, optimising with three objectives gives as good a convergence to the "true" two-objective Pareto-optimal fronts as when optimising with only two objectives. In other words, increasing the number of individuals in the population from 100 to 500 and the number of generations from 100 to 200 (with proportionately higher increase in processing time) ensures that the model adequately explores the 3D objective space.

[2] Recall that *LanPr* and *PolC* do not conflict; SO1SO3_LUPI gave a single optimum. This is why only SO3BO1BO2_LUPI-derived non-dominated set is set against three two-objective derived non-dominated fronts.

More specifically, it is capable of finding, amongst all compromise solutions, solutions that are, with respect only to the appropriate two objectives, equally good to those derived with scenarios where only two objectives are optimised.

Figures 8.32, 8.34 and 8.36, illustrate the 2D projections of the 3D surface-like shapes defined by the non-dominated sets obtained from the three-objective optimisations. By examining these graphs the second conclusion mentioned above becomes more obvious. Although in these cross-sections the majority of solutions appear to be dominated with respect to only the two objectives, the reader is reminded that all solutions are non-dominated to each other with respect to all three objectives.

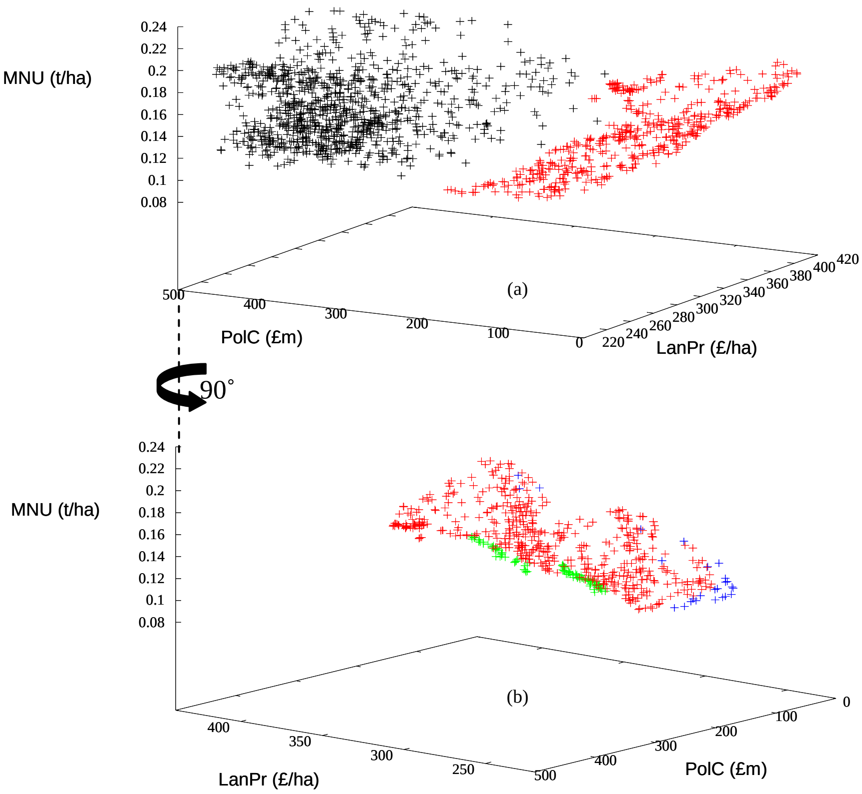

Figure 8.32. Three-objective optimisation results for SO1SO3BO1_LUPI. (a): non-dominated front (red) derived from optimising with SO1SO3BO1_LUPI in relation to the random initial set of solutions (black); (b): SO1SO3BO1_LUPI-derived non-dominated front together with the non-dominated fronts derived from optimising with SO1BO1_LUPI (green) and SO3BO1_LUPI (blue)

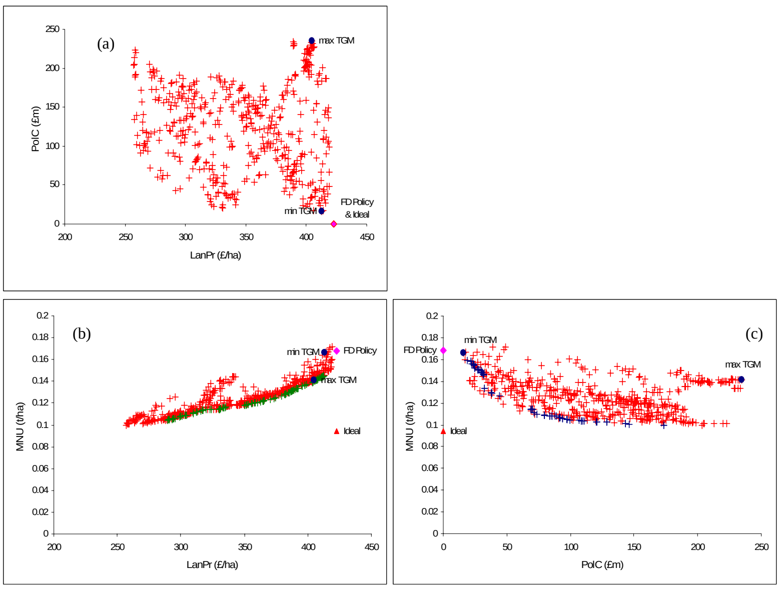

Figure 8.33. 2D projections of the SO1SO3BO1_LUPI-derived non-dominated front together with their two-objective counterparts. NOTE: the relationship between the cross-sections can be visualised by virtually rotating the shape in one cross-section in such a way as the min TGM and max TGM points of both cross-sections are finally matched.

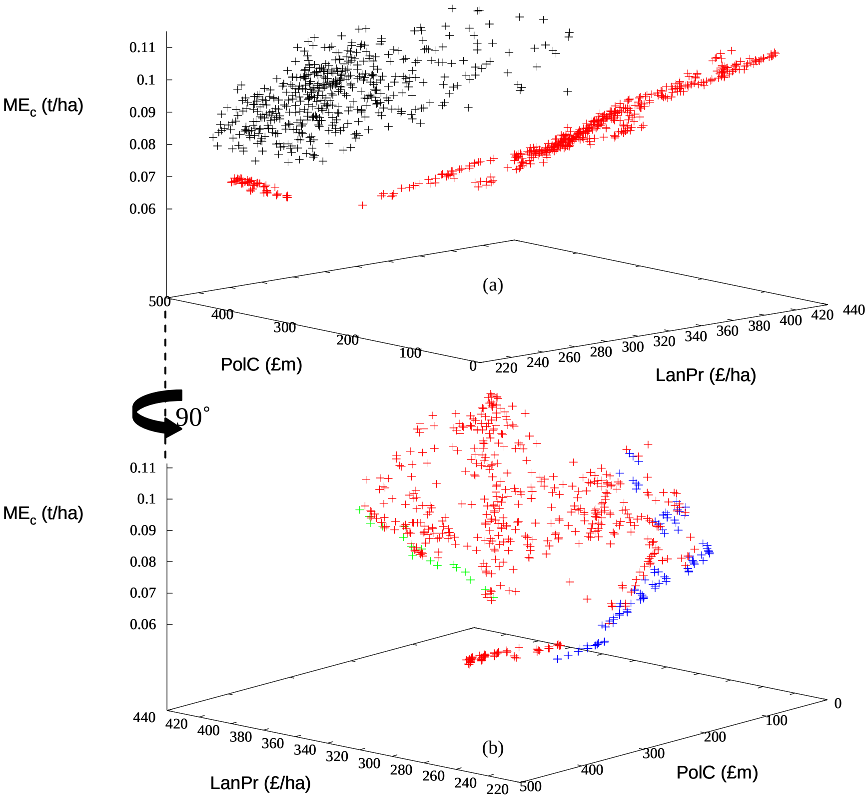

Figure 8.34. Three-objective optimisation results for SO1SO3BO2_LUPI. (a): non-dominated front (red) derived from optimising with SO1SO3BO2_LUPI in relation to the random initial set of solutions (black); (b): SO1SO3BO1_LUPI-derived non-dominated front together with the non-dominated fronts derived from optimising with SO1BO2_LUPI (green) & SO3BO2_LUPI (blue)

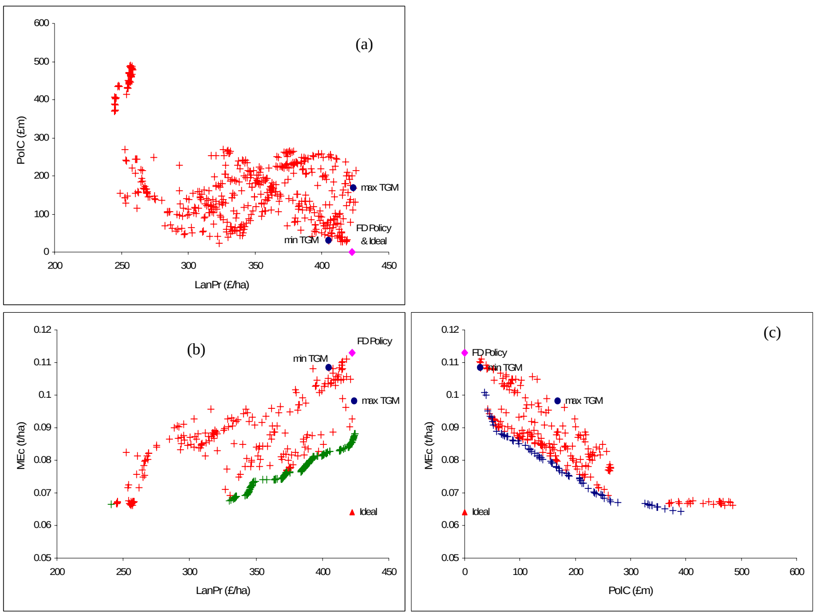

Figure 8.35. 2D projections of the SO1SO3BO2_LUPI-derived non-dominated front together with their two-objective counterparts. NOTE: the relationship between the cross-sections can be visualised by virtually rotating the shape in one cross-section in such a way as the min TGM and max TGM points of both cross-sections are finally matched.

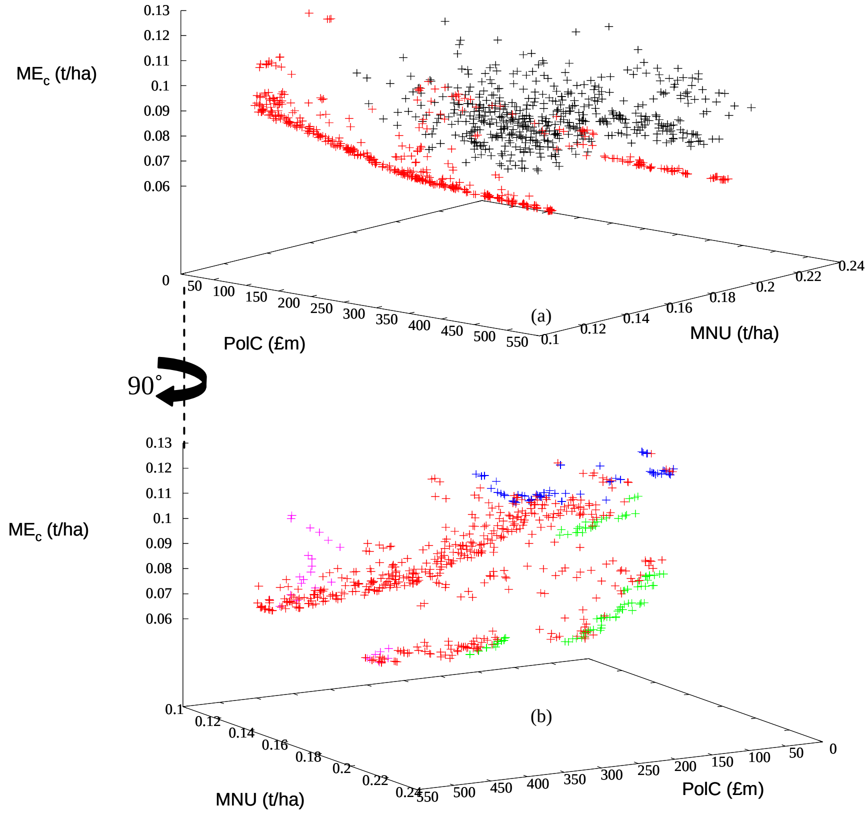

Figure 8.36. Three-objective optimisation results for SO3BO1BO2_LUPI. (a): non-dominated front (red) derived from optimising with SO3BO1BO2_LUPI together in relation to the random initial set of solutions (black); (b): SO3BO1BO2_LUPI-derived non-dominated front with the non-dominated fronts derived from optimising with BO1BO2_LUPI (purple), SO3BO1_LUPI (green) & SO3BO2_LUPI (blue).

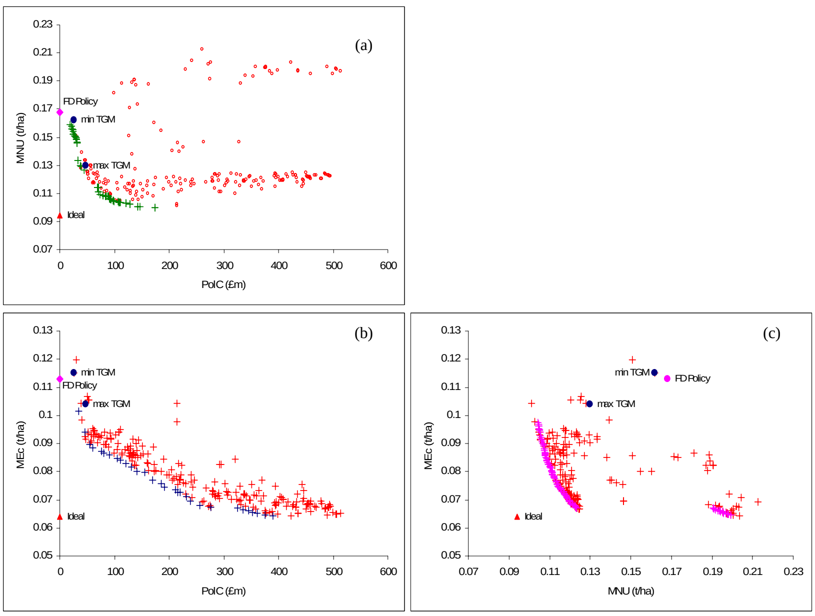

Figure 8.37. 2D projections of the SO3BO1BO2_LUPI-derived non-dominated front together with their two-objective counterparts. NOTE: the relationship between the cross-sections can be visualised by virtually rotating the shape in one cross-section in such a way as the min TGM and max TGM points of both cross-sections are finally matched.

8.4.3. Conclusion

Three scenarios of optimising agricultural policy with three objectives and with the LUPI policy instrument in place were examined. The structure of the resulting trade-off surfaces reveals the relationship between *LanPr*, *PolC*, *MNU* and ME_c. It is obvious that that the number of non-dominated solutions increases when optimising agricultural policy with three objectives. While this may be a desirable advantage in that it presents the DM with a large number of alternative policy options, at the same time it complicates the comparison of available options and effectively the decision making process. Storage and management of the model's output in database systems offers possibilities for handling and sorting large sets of non-dominated solutions. Running queries using case-specific criteria allows for the selection of the most suitable solution(s) whereby a final decision is more easily reached. Special multivariate data analysis techniques can be used for reducing the dimensionality of the output dataset and for exploring the association between the non-dominated solutions also with regard to the production activities and policy variables values. This can be particularly useful in situations where one is interested to investigate the relationship between objective achievement and policy variables and/or production decision variables. Thus, it can help address questions like: what production practices are associated positively or negatively with this or the other policy objective? Also, cluster analysis, for example, could be applied to form groups of policy solutions on the basis of their similarities using pre-specified criteria.

8.5. APPENDIX

Bar-Charts Legends

Cash Crops

- Set-Aside
- sWheat_160N
- sWheat_80N
- wBarley_180N
- wBarley-90N
- sBarley_100N
- sBarley_50N
- wOats_120N
- wOats_60N
- sOats_80N
- sOats_40N
- wOSR_185N
- wOSR_90N
- sOSR_110N
- sOSR_55N
- Triticale

Fodder Crops

- Rot_Grass_Grazing_125KgN
- Rot_Grass_Grazing_150KgN
- Rot_Cons_For_Hay_125KgN
- Rot_Cons_For_Hay_200KgN
- Rot_Cons_For_Sil_125KgN
- Rot_Cons_For_Sil_220KgN
- Turnips_Swedes_FodderBeet
- Rape
- Perm_Cons_For_Hay_125KgN
- Perm_Cons_For_Hay_200KgN
- Perm_Cons_For_Sil_125KgN
- Perm_Cons_For_Sil_220KgN
- Perm_Grass_Grazing_125KgN
- Perm_Grass_Grazing_175KgN

Cattle

1. Dcows_Spr-Sum_Calv
1a. Overwinter_Dcows_Spr-Sum_Calv
1b. Inten_Fin_DCalves
2. Dcows_Aut_Calv
3. Hill_SCows
3a. Overwinter_Hill_SCalves
3b. Inten_Fin_Hill_SCalves
4. Up/Low_Cattle_Fb-Jn_Sil
4a. Overwinter_Up/low_Calves_Fb-Jn_Sil
4b. Inten_Fin_Up/Low_Calves_Fb-Jn_Sil
5. Up/Low_Cattle_Ag-Oc_Sil
6. Up/Low_Cattle_Fb-Jn_Str
6a. Overwinter_Up/low_Calves_Fb-Jn_Str
6b. Inten_Fin_Up/Low_Calves_Fb-Jn_Str
7. Up/Low_Cattle_Ag-Oc_Str
8. Overwintering_SprCalves
9. Inten_Fin_Spr_Calves_12mon
10. Fin_AutCalves
11. Fin_SprCalves_18-24mon

Sheep

1. Hill_BreedEwes_StoreLamb
1a. Sell_Ewes_to_Up_for_Breed
1b. Store_Lambs
2. Hill_Ewes_for_FinLamb&GimProd
2a. Finish_Lambs
2b. Sell_for_Gim
2c. Store_Lambs
3. UpBreedEwes
3a. Sell_for_Gim
3b. Store_Lambs
3c. Finish_Lambs
4. LowBreedEwes_for_EarlyFinLamb
4a. Draft_Ewes
4b. Finish_Lambs
5. Draft_Ewes_for_Fin&StoreLambs
5a. Store_Lambs
5b. Finish_Lambs
6. Gimmerring
7. Fin_Lambs_Winter
8. Fin_Lambs_Autumn

Table 8.5.1. Policy Variables Values for SO1BO1_LUPI (1)

code	Land Pr	MNU	Set-Aside	sWheat_160N	sWheat_80N	wBarley_180N	wBarley_90N	sBarley_100N	sBarley_50N	wOats_120N	wOats_60N	sOats_80N	sOats_40N	wOSR_185N	wOSR_90N	sOSR_110N	sOSR_55N	Triticale
70	290.4	0.105	300	0	300	180	240	0	300	0	260	60	280	20	160	20	160	20
52	295.6	0.106	220	100	280	20	300	20	300	0	240	140	240	100	240	60	240	20
39	300.1	0.108	180	20	280	20	260	140	300	160	220	140	300	20	180	220	20	80
68	304.6	0.109	120	0	300	0	300	80	300	200	200	80	300	0	140	280	200	280
88	309.8	0.110	80	20	300	80	300	180	300	100	300	80	300	60	280	220	280	60
85	315.5	0.113	80	0	220	100	300	140	300	260	300	0	280	20	300	200	20	60
20	320.1	0.114	300	160	280	120	300	100	300	20	260	120	240	0	260	160	200	240
90	324.2	0.114	300	0	300	240	300	120	300	80	220	40	280	20	80	180	280	160
34	330.1	0.114	300	0	300	240	300	120	300	0	220	40	280	20	280	180	280	160
16	335.0	0.116	280	20	260	100	280	0	300	20	100	40	300	20	100	220	160	80
19	355.7	0.119	260	0	220	180	260	60	300	20	160	120	280	20	80	220	60	240
62	360.1	0.120	220	200	300	140	300	180	300	100	300	80	300	20	120	60	280	60
69	365.1	0.122	180	0	300	240	300	20	300	180	220	80	280	0	120	180	180	220
100	370.6	0.124	100	0	280	120	300	40	300	140	220	80	140	20	200	220	40	60
83	374.4	0.125	80	0	300	140	300	180	300	100	220	80	260	20	280	220	280	20
10	381.8	0.128	20	100	300	180	300	100	300	40	200	80	160	40	100	20	80	100
43	385.1	0.129	40	80	300	120	300	80	300	100	300	40	280	80	220	240	140	180
55	390.2	0.133	40	0	300	120	300	80	300	160	300	40	300	80	220	260	300	180
14	395.2	0.136	60	60	300	100	280	120	280	60	160	180	60	60	120	20	40	280
41	399.2	0.137	20	20	300	120	300	180	280	180	220	100	300	0	80	200	80	140
17	405.5	0.140	40	160	300	40	300	180	300	180	220	80	280	20	80	300	120	60
12	410.0	0.143	140	20	300	100	280	60	220	100	300	280	120	0	80	40	60	60

Table 8.5.2. Policy Variables Values for SO1BO1_LUPI (2)

code	LanPr	MNU	Rot_Grass_Grazing_125KgN	Rot_Grass_Grazing_150KgN	Rot_Cons_For_Hay_125KgN	Rot_Cons_For_Hay_200KgN	Rot_Cons_For_Sil_125KgN	Rot_Cons_For_Sil_220KgN	Turnips_Swedes_FodderBeet	Rape	Perm_Cons_For_Hay_125KgN	Perm_Cons_For_Hay_200KgN	Perm_Cons_For_Sil_125KgN	Perm_Cons_For_Sil_220KgN	Perm_Grass_Grazing_125KgN	Perm_Grass_Grazing_175KgN
70	290.4	0.105	0	0	60	60	0	0	20	40	20	40	100	80	0	60
52	295.6	0.106	0	0	0	40	0	20	20	40	20	40	60	0	20	260
39	300.1	0.108	0	0	40	40	0	20	0	60	0	60	120	0	20	220
68	304.6	0.109	0	0	0	40	0	20	20	40	20	60	60	0	0	280
88	309.8	0.110	0	0	0	40	0	20	20	40	60	60	60	0	0	280
85	315.5	0.113	0	0	60	60	0	0	20	40	40	40	100	0	0	300
20	320.1	0.114	20	0	40	40	100	120	0	0	140	60	0	20	20	280
90	324.2	0.114	20	0	40	40	0	20	20	0	100	40	0	0	20	260
34	330.1	0.114	20	0	0	40	0	0	20	0	20	40	40	0	0	140
16	335.0	0.116	20	0	20	40	80	0	0	0	80	20	0	40	0	280
19	355.7	0.119	0	0	0	40	0	20	20	0	20	40	40	0	20	260
62	360.1	0.120	0	0	0	40	0	20	20	0	0	60	40	0	40	220
69	365.1	0.122	0	0	40	40	0	20	20	0	0	60	60	20	0	120
100	370.6	0.124	0	0	0	20	0	0	20	0	20	40	20	20	20	240
83	374.4	0.125	0	0	0	20	0	20	40	0	0	40	20	20	20	220
10	381.8	0.128	0	0	20	60	20	40	20	0	80	20	80	40	0	280
43	385.1	0.129	0	0	20	60	0	20	100	60	80	0	80	0	0	300
55	390.2	0.133	0	0	20	60	0	0	260	40	20	60	60	0	0	140
14	395.2	0.136	0	0	20	80	20	40	260	0	120	40	80	40	0	280
41	399.2	0.137	0	0	120	200	0	20	20	0	20	120	220	40	0	200
17	405.5	0.140	0	0	100	240	20	40	0	40	220	140	300	140	40	220
12	410.0	0.143	0	0	60	200	20	40	0	40	200	120	300	80	40	140

Table 8.5.3. Policy Variables Values for SO1BO2_LUPI (1)

code	LanPr	MEc	Set-Aside	sWheat_160N	sWheat_80N	wBarley_180N	wBarley_90N	sBarley_100N	sBarley_50N	wOats_120N	wOats_60N	sOats_80N	sOats_40N	wOSR_185N	wOSR_90N	sOSR_110N	sOSR_55N	Triticale
87	325.0	0.068	80	0	0	0	20	0	40	180	160	0	20	0	20	100	80	40
64	335.1	0.069	300	0	40	0	20	0	60	80	60	0	20	0	40	100	0	0
12	345.5	0.071	260	0	20	0	0	0	60	80	20	0	0	0	60	100	100	40
34	349.8	0.073	40	0	20	40	80	0	20	80	80	0	20	0	0	140	120	0
78	355.5	0.074	260	0	40	0	0	0	60	100	160	140	60	20	60	0	120	240
65	359.8	0.074	280	0	40	0	0	0	40	120	160	0	0	0	20	100	80	180
44	364.0	0.074	260	0	40	0	0	0	60	80	0	0	20	0	20	20	80	0
15	370.0	0.075	300	0	40	0	0	0	60	80	60	0	20	0	60	100	80	0
84	376.9	0.076	280	0	40	0	0	0	40	180	160	0	20	0	60	20	80	20
93	381.5	0.076	240	0	40	0	0	0	0	80	180	0	20	0	40	60	60	40
96	385.1	0.077	240	0	40	0	0	0	60	120	60	0	20	0	0	140	100	40
67	390.4	0.079	20	0	0	60	20	0	40	300	160	120	20	20	40	100	160	180
88	395.1	0.082	240	0	40	0	60	0	60	80	20	0	20	0	0	60	20	0
51	398.2	0.083	300	0	40	0	0	0	60	80	40	0	20	0	20	100	140	60
55	406.8	0.083	260	0	40	0	0	0	60	80	0	0	20	0	0	100	240	60
20	414.0	0.083	280	0	40	0	40	0	20	80	0	0	20	0	40	100	80	120
71	420.8	0.085	0	0	0	0	0	0	40	40	80	0	20	0	0	80	140	160

Table 8.5.4. Policy Variables Values for SO1BO2_LUPI (2)

code	LanPr	MEc	Rot_Grass_Grazing_125KgN	Rot_Grass_Grazing_150KgN	Rot_Cons_For_Hay_125KgN	Rot_Cons_For_Hay_200KgN	Rot_Cons_For_Sil_125KgN	Rot_Cons_For_Sil_220KgN	Turnips_Swedes_FodderBeet	Rape	Perm_Cons_For_Hay_125KgN	Perm_Cons_For_Hay_200KgN	Perm_Cons_For_Sil_125KgN	Perm_Cons_For_Sil_220KgN	Perm_Grass_Grazing_125KgN	Perm_Grass_Grazing_175KgN
87	325.0	0.068	20	20	300	220	300	140	0	0	240	260	300	100	40	260
64	335.1	0.069	0	0	300	220	120	80	0	20	300	200	300	140	60	220
12	345.5	0.071	0	0	300	220	0	20	0	20	300	0	300	120	300	120
34	349.8	0.073	0	20	300	280	0	0	20	20	280	200	300	140	60	300
78	355.5	0.074	0	60	300	300	0	20	40	20	300	200	300	140	240	40
65	359.8	0.074	0	20	300	300	40	60	0	0	280	160	280	240	280	0
44	364.0	0.074	0	20	300	300	300	0	0	20	20	60	280	0	120	80
15	370.0	0.075	0	20	300	300	180	100	0	0	220	260	300	140	80	0
84	376.9	0.076	20	0	300	300	300	40	0	20	20	40	300	120	300	280
93	381.5	0.076	0	40	300	300	0	20	0	20	260	160	280	120	120	180
96	385.1	0.077	20	40	300	300	40	20	20	20	300	240	300	160	300	280
67	390.4	0.079	0	0	300	300	0	60	40	20	300	200	300	140	280	20
88	395.1	0.082	0	20	280	280	0	0	20	0	280	160	300	140	60	280
51	398.2	0.083	0	0	0	300	300	0	40	40	160	20	300	20	280	120
55	406.8	0.083	0	80	180	300	300	300	40	20	260	160	300	40	240	0
20	414.0	0.083	0	0	120	300	40	60	20	20	300	220	300	80	300	280
71	420.8	0.085	20	80	80	300	20	0	20	20	300	280	300	60	60	220

Table 8.5.5. Policy Variables Values for SO3BO1_LUPI (1)

code	PolC	MNU	Set-Aside	sWheat_160N	sWheat_80N	wBarley_180N	wBarley_90N	sBarley_100N	sBarley_50N	wOats_120N	wOats_60N	sOats_80N	sOats_40N	wOSR_185N	wOSR_90N	sOSR_110N	sOSR_55N	Triticale
73	20.6	0.158	40	20	80	0	80	0	40	0	160	40	20	60	80	120	100	20
91	25.2	0.151	120	20	80	40	100	0	40	80	160	40	20	20	80	40	40	20
18	30.0	0.148	40	20	80	40	100	0	40	80	160	40	20	60	80	120	140	20
81	34.2	0.131	40	20	80	40	100	0	0	80	160	60	100	20	100	120	40	20
59	41.0	0.129	40	20	80	40	20	0	40	0	0	40	20	220	120	120	60	20
80	46.1	0.115	100	20	80	0	100	0	0	80	180	0	20	60	120	120	140	20
66	64.1	0.115	140	20	80	0	100	0	0	80	160	60	100	20	260	40	120	20
6	70.0	0.114	140	20	80	40	100	0	80	80	160	40	100	20	100	40	40	20
92	74.1	0.109	120	0	80	40	100	0	80	0	200	40	100	0	80	40	40	20
85	82.5	0.108	140	20	80	0	100	0	80	80	200	60	100	0	120	100	220	0
45	85.3	0.108	140	20	80	40	100	0	80	80	160	40	220	0	120	100	300	0
36	91.3	0.106	220	0	80	40	100	0	80	0	200	40	100	0	80	40	40	20
29	94.0	0.106	240	20	80	0	100	0	80	80	160	40	80	20	80	40	40	20
95	99.6	0.104	300	20	80	20	100	0	80	80	160	40	100	20	100	40	40	20
57	104.9	0.104	300	20	80	0	100	0	80	80	180	120	220	0	60	80	300	20
39	110.6	0.104	300	20	80	0	100	0	80	80	180	120	220	0	60	80	300	20
33	120.2	0.103	300	20	80	40	100	0	120	140	200	120	220	0	120	80	40	100
23	125.1	0.103	300	20	120	40	100	0	120	140	200	120	220	0	120	100	300	0
54	129.9	0.102	300	20	80	60	160	40	160	100	240	40	260	80	140	80	280	20
53	136.5	0.102	300	20	80	60	160	40	160	100	240	40	260	0	140	80	280	20
43	142.8	0.101	300	0	60	100	160	120	220	0	220	120	220	0	60	80	300	20
76	146.2	0.100	300	0	60	100	160	120	220	0	220	80	260	80	140	80	280	20
82	153.4	0.100	300	20	80	0	240	120	220	0	220	80	260	80	140	80	280	20
4	173.6	0.100	300	20	80	40	100	20	280	80	200	120	180	0	140	80	300	20
88	182.8	0.099	300	20	80	80	200	140	300	80	240	80	260	80	140	80	300	20

Table 8.5.6. Policy Variables Values for SO3BO1_LUPI (2)

code	PolC	MNU	Rot_Grass_Grazing_125KgN	Rot_Grass_Grazing_150KgN	Rot_Cons_For_Hay_125KgN	Rot_Cons_For_Hay_200KgN	Rot_Cons_For_Sil_125KgN	Rot_Cons_For_Sil_220KgN	Turnips_Swedes_FodderBeet	Rape	Perm_Cons_For_Hay_125KgN	Perm_Cons_For_Hay_200KgN	Perm_Cons_For_Sil_125KgN	Perm_Cons_For_Sil_220KgN	Perm_Grass_Grazing_125KgN	Perm_Grass_Grazing_175KgN
73	20.6	0.158	0	0	0	0	180	40	0	0	20	20	0	20	0	0
91	25.2	0.151	0	0	0	0	20	120	0	0	180	20	0	20	0	0
18	30.0	0.148	0	0	0	20	180	40	0	0	20	60	0	20	0	0
81	34.2	0.131	0	0	0	0	20	40	0	40	20	20	0	20	0	60
59	41.0	0.129	20	0	0	0	20	120	0	20	180	20	0	20	0	20
80	46.1	0.115	20	0	20	0	20	40	0	40	20	0	0	20	0	20
66	64.1	0.115	40	0	0	0	20	40	0	40	20	20	0	20	0	60
6	70.0	0.114	0	0	0	20	180	120	0	40	180	60	0	20	0	80
92	74.1	0.109	40	20	40	0	20	40	0	40	20	20	0	20	0	80
85	82.5	0.108	40	0	40	20	100	120	0	40	100	20	0	20	0	0
45	85.3	0.108	40	0	40	20	100	120	0	40	20	20	0	20	0	0
36	91.3	0.106	40	20	40	0	100	40	0	40	20	20	0	20	0	80
29	94.0	0.106	40	0	40	0	100	40	0	40	20	20	0	20	0	0
95	99.6	0.104	40	0	0	0	100	40	0	40	20	20	0	20	0	60
57	104.9	0.104	40	0	0	20	100	120	0	40	0	20	0	20	0	0
39	110.6	0.104	40	0	40	20	100	120	0	40	20	20	0	20	0	20
33	120.2	0.103	40	0	40	0	100	40	0	40	20	20	0	20	0	0
23	125.1	0.103	40	0	40	20	100	120	0	40	20	20	0	20	0	0
54	129.9	0.102	40	0	0	0	100	40	0	40	20	20	0	20	0	60
53	136.5	0.102	40	0	40	0	80	40	0	40	20	60	0	20	0	20
43	142.8	0.101	40	0	0	0	60	80	0	40	20	20	0	20	0	0
76	146.2	0.100	40	0	0	0	20	40	0	40	20	20	0	20	40	60
82	153.4	0.100	40	0	0	0	20	40	0	40	20	20	0	20	40	60
4	173.6	0.100	40	0	0	0	20	120	0	40	180	20	0	20	0	0

Table 8.5.7. Policy Variables Values for BO1BO2_LUPI (1)

code	MNU	MEc	Set-Aside	sWheat_160N	sWheat_80N	wBarley_180N	wBarley_90N	sBarley_100N	sBarley_50N	wOats_120N	wOats_60N	sOats_80N	sOats_40N	wOSR_185N	wOSR_90N	sOSR_110N	sOSR_55N	Triticale
16	0.104	0.097	300	20	160	0	100	220	300	60	220	0	140	80	260	100	260	200
11	0.105	0.096	300	20	160	0	100	60	300	60	220	0	140	80	180	100	260	200
91	0.106	0.092	300	20	80	0	100	0	240	20	300	80	240	0	100	200	140	40
70	0.107	0.09	300	20	80	0	80	20	200	80	140	20	140	80	180	120	300	40
25	0.108	0.087	300	0	80	0	80	0	140	80	140	80	160	0	140	180	300	60
68	0.109	0.084	300	0	80	0	100	0	80	20	160	40	120	20	80	120	220	40
40	0.11	0.082	300	20	80	0	80	0	80	100	160	0	160	80	180	80	240	220
77	0.111	0.081	300	0	80	0	80	0	100	20	240	200	240	0	80	80	300	40
56	0.112	0.079	300	20	80	0	80	0	80	60	160	80	160	0	60	180	80	0
71	0.113	0.077	300	0	80	0	80	0	80	20	160	80	160	0	120	180	0	0
2	0.114	0.076	300	0	80	0	80	0	80	20	240	40	280	20	100	120	300	40
43	0.115	0.075	300	20	80	0	80	0	80	20	260	20	80	0	20	60	160	60
58	0.116	0.074	300	0	80	0	80	0	80	20	160	40	160	0	60	20	40	200
45	0.117	0.073	300	20	80	0	80	0	80	20	240	40	280	20	100	80	300	60
20	0.118	0.072	300	20	80	0	80	0	80	40	300	0	140	20	100	40	300	60
52	0.119	0.07	300	20	80	0	80	0	80	20	240	0	160	0	60	180	0	40
32	0.12	0.07	300	20	80	0	80	0	80	20	260	20	80	0	20	20	160	60
62	0.121	0.069	300	0	80	0	80	0	40	20	20	80	160	0	20	20	0	40
61	0.123	0.068	300	0	80	0	80	0	80	20	180	80	0	0	20	20	0	40
73	0.19	0.067	300	20	80	0	80	0	80	0	260	0	160	0	60	180	40	40
97	0.192	0.067	300	20	80	20	80	0	80	0	260	0	160	0	60	180	40	40
63	0.194	0.066	300	0	80	0	80	0	80	100	160	40	280	20	80	240	240	40
42	0.195	0.066	300	20	80	0	80	0	80	80	260	20	160	40	100	160	260	60
96	0.196	0.065	300	20	80	0	80	0	80	20	240	20	80	0	20	20	160	60
27	0.197	0.065	300	20	80	0	80	0	40	20	80	0	0	0	20	20	0	40
64	0.199	0.065	300	0	0	0	80	0	40	0	180	80	160	0	0	100	0	20

Table 8.5.8. Policy Variables Values for BO1BO2_LUPI (2)

code	MNU	MEc	Rot_Grass_Grazing_125KgN	Rot_Grass_Grazing_150KgN	Rot_Cons_For_Hay_125KgN	Rot_Cons_For_Hay_200KgN	Rot_Cons_For_Sil_125KgN	Rot_Cons_For_Sil_220KgN	Turnips_Swedes_FodderBeet	Rape	Perm_Cons_For_Hay_125KgN	Perm_Cons_For_Hay_200KgN	Perm_Cons_For_Sil_125KgN	Perm_Cons_For_Sil_220KgN	Perm_Grass_Grazing_125KgN	Perm_Grass_Grazing_175KgN
16	0.104	0.097	280	120	160	60	300	160	140	200	160	80	0	60	40	0
11	0.105	0.096	280	120	160	60	300	160	80	200	240	80	0	60	20	40
91	0.106	0.092	280	120	160	60	280	200	100	200	240	80	0	60	20	0
70	0.107	0.09	280	120	160	60	300	160	100	200	240	80	0	60	20	40
25	0.108	0.087	280	120	160	60	300	0	100	200	280	80	0	60	60	0
68	0.109	0.084	280	120	160	140	300	160	100	200	240	80	0	60	20	0
40	0.11	0.082	300	200	180	140	140	20	160	240	300	80	80	20	140	300
77	0.111	0.081	300	120	220	200	140	160	100	200	280	80	0	60	60	0
56	0.112	0.079	300	80	240	40	120	160	20	200	240	80	0	60	60	0
71	0.113	0.077	300	80	280	260	220	160	100	160	280	120	0	60	20	0
2	0.114	0.076	300	280	300	40	220	160	100	160	280	80	100	60	260	80
43	0.115	0.075	300	280	300	0	220	0	60	160	280	160	140	60	180	0
58	0.116	0.074	240	120	300	0	220	40	60	200	280	0	100	60	180	280
45	0.117	0.073	140	100	300	220	260	160	100	160	300	80	160	80	300	0
20	0.118	0.072	120	140	300	220	260	160	120	200	240	160	160	300	100	300
32	0.12	0.07	120	140	300	220	260	160	140	280	240	160	120	80	300	0
62	0.121	0.069	80	20	300	300	300	160	300	240	240	160	120	0	300	0
61	0.123	0.068	0	20	300	220	300	160	300	280	280	160	120	80	300	0
73	0.19	0.067	300	180	180	140	180	60	140	0	300	240	160	240	140	100
97	0.192	0.067	300	180	180	140	180	60	140	0	300	240	160	260	300	120
63	0.194	0.066	300	120	260	100	160	140	60	0	280	160	180	280	20	300
42	0.195	0.066	300	80	160	0	180	60	60	120	300	160	160	300	20	300
96	0.196	0.065	300	80	260	100	160	140	60	0	280	160	180	300	60	0
27	0.197	0.065	300	240	260	100	160	140	60	0	240	160	180	280	20	20
64	0.199	0.065	300	100	240	80	180	60	60	40	280	160	160	300	100	0

Table 8.5.9. Policy Variables Values for SO1BO1_RI

	Policy Objectives		Policy Constraints to Agricultural Production			
code	LanPr (£/ha)	MNU (t/ha)	MSA	QFC	TE_c	TNU
11	310.76	0.107	265396	457771	143695	100293
25	321.91	0.111	236950	310948	160117	100147
8	324.56	0.111	230499	447410	157771	100147
5	326.17	0.112	226393	461095	155572	100000
17	328.89	0.113	219941	453275	198680	100000
21	343.87	0.117	185924	459726	199267	100000
10	360.50	0.122	151320	348876	157771	100147
15	364.39	0.124	143988	378397	135484	100440
29	366.18	0.124	139883	474585	157771	100000
7	370.89	0.126	131378	496676	161290	100440
18	372.74	0.126	127273	500000	132258	100000
28	379.03	0.128	115543	433529	167889	100000
27	385.52	0.130	103812	477126	195161	100000
14	387.67	0.130	100000	329717	123900	100000
23	390.18	0.131	95601	459140	175513	100000
2	390.85	0.131	94428	459140	177859	100000
1	399.64	0.134	79472	389736	138856	100000
9	402.64	0.135	74487	347312	138123	100000
20	404.43	0.135	71554	347312	157771	100000
4	408.59	0.137	64809	474585	143402	100000
30	410.24	0.137	62170	447410	157331	100000
26	412.65	0.138	58358	415934	179765	100000
13	414.89	0.138	54839	349658	218768	100000
3	422.34	0.141	43402	343011	123460	100000
22	426.65	0.142	36950	373118	160704	100000
19	429.95	0.143	32258	464614	216422	100147
24	431.84	0.144	29326	439198	146041	100000
12	434.40	0.144	25806	301760	128886	100147
6	443.13	0.147	13490	391496	151320	100147
16	446.02	0.148	9677	349853	246774	100293

Table 8.5.10. Policy Variables Values for SO1BO2_RI

	Policy Objectives		**Policy Constraints to Agricultural Production**			
code	LanPr (£/ha)	ME_c (t/ha)	MSA	QFC	TE_c	TNU
5	325.37	0.093	240762	490029	100000	217595
25	328.62	0.093	231378	490029	100000	199413
2	335.16	0.094	212903	409091	100000	178299
7	335.79	0.095	211144	496285	100000	177126
17	339.40	0.095	201173	427859	100000	213490
12	342.19	0.096	193548	477517	100000	231672
3	347.65	0.097	178886	402835	100000	180059
18	351.20	0.097	169501	446628	100000	223021
15	353.56	0.098	163343	487683	100000	218182
24	356.85	0.098	154839	477908	100000	182991
28	362.16	0.099	141349	480645	100000	232845
6	362.39	0.099	140762	493157	100000	232258
26	365.91	0.100	131965	416520	100000	199413
8	368.05	0.101	126686	488074	100000	197067
11	370.17	0.101	121408	439003	100147	187537
4	373.82	0.102	112610	452884	100000	180059
22	373.82	0.102	112610	452493	100000	234018
10	377.50	0.102	103812	394233	100000	207625
14	377.50	0.102	103812	394233	100000	224927
30	381.8424	0.103056	93548	427859	100000	175367
9	386.89	0.104	81818	455621	100000	217742
23	389.84	0.105	75073	406158	100000	222727
21	395.30	0.106	62757	459335	100000	206452
1	397.94	0.106	56891	387586	100000	213490
13	401.41	0.107	49267	395406	100000	219355
16	401.41	0.107	49267	395406	100000	207625
19	406.00	0.108	39296	494330	100000	180059
20	408.74	0.108	33431	415347	100000	206452
29	423.83	0.111	2053	490029	100000	217595
27	423.98	0.111	1760	452884	100000	217595

Table 8.5.11. Policy Variables Values for SO2BO1_RI

	Policy Objectives		Policy Constraints to Agricultural Production			
Code	LabPr (£/ha)	MNU (t/ha)	MSA	QFC	TE_c	TNU
28	10.003	0.153	73607	314272	125513	213783
25	9.971	0.138	175953	314663	206745	138856
13	9.971	0.138	175953	314272	211437	138856
3	9.956	0.136	187390	314663	146041	138856
15	9.941	0.135	196774	314272	126686	138856
7	9.928	0.134	204985	314076	166569	138710
6	9.916	0.133	213783	314272	248974	138856
16	9.899	0.132	223167	314272	211437	139003
1	9.886	0.130	234311	314272	126686	138856
5	9.873	0.129	243695	314272	146041	138856
4	9.852	0.128	253079	314272	129032	138710
2	9.827	0.127	130499	416911	243255	101026
20	9.816	0.125	281232	314272	168182	139003
14	9.796	0.124	151613	414174	210850	101466
9	9.791	0.123	298827	314272	125513	139003
29	9.771	0.122	151906	338905	247947	100000
26	9.771	0.122	156892	313099	216129	100440
8	9.739	0.119	175660	314663	214956	100440
10	9.721	0.117	196774	318377	206745	101466
22	9.705	0.116	204985	312512	165982	101320
11	9.683	0.115	204985	341251	173167	100147
24	9.677	0.114	223167	339296	211144	101466
12	9.656	0.112	233724	314272	102346	101613
23	9.643	0.112	227273	338905	243109	100000
27	9.603	0.109	252493	441349	173167	100147
17	9.597	0.108	270675	364125	210850	101466
30	9.580	0.107	280645	464418	163636	101466
18	9.579	0.107	281232	364321	158651	101466
19	9.562	0.106	290616	314272	102346	101613
21	9.542	0.105	290323	415152	205718	100293

Table 8.5.12. Policy Variables Values for SO2BO2_RI

	Policy Objectives		Policy Constraints to Agricultural Production			
Code	LabPr (£/ha)	MNU (t/ha)	MSA	QFC	TE_c	TNU
1	9.82	0.0968	246334.3	425904	102639	137683
2	9.86	0.0980	247507.3	416325	104106	137537
3	9.90	0.0991	225513.2	402053	103079	137537
4	9.75	0.0949	254545.5	351808	101173	137830
5	9.91	0.0997	194721.4	468328	100587	137683
6	9.84	0.0976	219061.6	416129	100733	137683
7	9.63	0.0917	285630.5	480841	100587	138416
8	9.70	0.0934	271847.5	422385	101026	137683
9	9.72	0.0939	287096.8	393842	103079	137537
10	9.79	0.0959	285044	491789	105572	137830
11	9.66	0.0925	283871	425904	101026	137683
12	9.80	0.0962	251319.6	466569	102346	137537
13	9.73	0.0943	281818.2	400489	103079	137683
14	9.88	0.0986	219061.6	463050	101906	137683
15	9.67	0.0928	271554.3	477126	100293	137830
16	9.94	0.1002	204398.8	403226	102199	137537
17	9.61	0.0912	299413.5	422385	100880	137683
18	9.69	0.0931	299120.3	457380	103372	137683
19	9.76	0.0952	284457.5	490225	104399	137537
20	9.83	0.0972	227272.7	376442	101173	137830
21	9.60	0.0910	292961.9	353568	100733	138856
22	9.65	0.0920	298827	468915	101906	137683
23	9.74	0.0946	248387.1	452102	100000	137537
24	9.84	0.0974	254252.2	455230	104106	137537
25	9.57	0.0902	299413.5	416129	100293	138856
26	9.59	0.0906	299413.5	422385	100587	138416
27	9.98	0.1031	155718.5	393842	103079	142229
28	9.70	0.0938	265102.6	403226	101320	138563
29	9.77	0.0955	255425.2	488074	101906	137683
30	9.93	0.1001	204398.8	428250	102053	137537

Table 8.5.13. Policy Variables Values for BO1BO2_RI

	Policy Objectives		Policy Constraints to Agricultural Production			
Code	MNU (t/ha)	ME_c (t/ha)	MSA	QFC	TE_c	TNU
3	0.1032	0.1029	300000	494917	180059	100000
25	0.1036	0.1029	300000	496872	100000	100440
22	0.1058	0.1028	299414	321114	100733	103666
2	0.1062	0.1024	300000	484360	100000	103666
16	0.1092	0.1023	300000	312708	100293	107038
29	0.1092	0.1023	300000	300196	100293	107038
26	0.1098	0.1013	300000	312512	100000	108358
23	0.1102	0.1010	300000	435484	100147	109238
11	0.1109	0.1010	300000	482796	100000	109824
7	0.1117	0.1002	300000	487879	100293	111730
24	0.1123	0.0992	300000	338710	100147	113343
4	0.1123	0.0991	300000	301173	100000	113343
27	0.1133	0.0986	300000	312708	101320	116422
15	0.1134	0.0985	300000	312708	101320	116716
8	0.1144	0.0972	300000	312512	100000	117742
13	0.1154	0.0968	300000	484360	100000	119208
9	0.1167	0.0963	300000	437634	100293	121554
21	0.1179	0.0961	300000	313490	101026	123900
14	0.1186	0.0954	299414	487879	100000	124340
19	0.1198	0.0952	300000	313490	100880	126979
12	0.1208	0.0944	300000	332649	100000	128006
5	0.1218	0.0938	299707	301173	100733	130792
18	0.1221	0.0930	300000	300196	100293	137830
28	0.1230	0.0924	300000	490420	100000	133138
10	0.1230	0.0918	300000	312512	100000	188856
6	0.1236	0.0911	300000	321114	100000	187390
17	0.1240	0.0907	300000	326002	100000	226100
30	0.1251	0.0897	299414	340274	100000	170674
1	0.1256	0.0890	300000	487879	100000	187390
20	0.1256	0.0890	300000	440371	100000	170674

9. Summary and Conclusions

9.1. Thesis Overview

The problem of optimisation of agricultural policy comprised the thematic core of the research presented in this contribution. The study was focused on two interrelated subproblems of policy analysis: the follower's problem, concerned with the prediction of farmers' responses to the policy actions; and the leader's problem concerned with the configuration of optimal policies. The problem has been formulated as multi-objective bi-level programming whereby the two subproblems are integrated directly into one policy optimisation model, taking into account the hierarchical nature of the decision-making process. The general formulation of the model described by equations (1a)-(2d) is a nonlinear bi-level and multi-objective programming problem hard to solve. Complicating nonlinearities appear in situations where a leader's objective function contains cross-products of leader's and follower's decision variables (in the application model (7.11) the cross-product term, $S_j \cdot X_j$ (area allocated to crop *j* x payment for crop) in the *PolC* objective function for example). The additional fact that the follower's constraint region is dependent on the leader's regulatory constraint variables and the existence of multiple objective functions further complicates the solution process.

Today non-heuristic algorithms for solving multi-objective bi-level programming problem are, to the best of the author's knowledge, non existent. A small number of computer codes exist for classical algorithms for solving linear single-objective bi-level programming problems. At best, they can handle problems of size limited to 200-300 leader variables, 100-150 follower variables and 50 constraints. When nonlinearities are present, the manageable size shrinks by nearly an order of magnitude (Bard, 1998). For the Scottish case-study presented here the leader sub-problem has up to three objective functions, 30 variables and two constraints whereas, the follower sub-problem has on objective function, 88 variables and 45 constraints. Therefore, classical bi-level programming algorithms would not be able to handle the present problem and a different approach had to be developed.

Recall that in the formulation of the model as a BLP problem, the government is given control over the policy decision variables, while the agricultural sector collectively controls the production decision variables. The policy maker makes a guess and anticipates the response of the farmers before selecting a policy. The approach proposed exploits the fact that once the policy variables are specified, the follower's model is parameterised in the leader's variables and reduces to a quadratic programming problem which can be readily solved with available quadratic programming solvers. In addition, it takes the view that policy optimisation can be visualised as an evolutionary process of continuous improvement subject to a set of conditions partly determined by the farmers' reactions.

A genetic algorithm can simulate in a logical and natural way the above procedures by using operators that mimic the biological processes of natural evolution. With special modifications genetic algorithms can handle multiple objective optimisation problems. In addition, genetic algorithms can deal successfully with non-smooth, non-continuous and non-differentiable functions. Thus, the approach avoids the difficulties resulting from the non-differentiability of the mathematical relationship between the leader's and follower's decision variables. Additionally, modularisation of the optimisation domains greatly increases the approach's efficiency. Overall, the hybridisation of a genetic algorithm with classical optimisation methods results in powerful optimisation process suited for multi-objective bi-level programming problems. Figure 9.1 illustrates a reminder of the stages involved in the hybrid algorithm developed for the optimisation of agricultural policy in Scotland.

The model for the follower's problem is probably the most essential part of the APOLO model cycle in that it determines the degree of realism in forecasting the farmer's responses to exogenous changes. Hence, the reliability of the results produced by APOLO depends a great deal on the performance of the SASEM model developed for the application. The generalised maximum entropy procedure, within a positive mathematical programming framework, was used for the estimation of the farmer's model objective function parameters. The positive nature of the resulting model enhances the confidence on the model's predictions of farmers' collective responses to changes in exogenous variables such as policy measures. Based on observations from three years, reliable estimations for the cost function parameters were obtained. The estimation of all the off-diagonal elements of the variable costs matrix recovers important information about complementarity and substitution relations between activities.

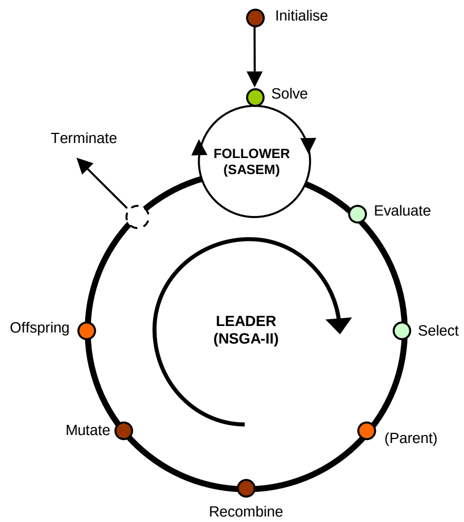

Figure 9.1. The APOLO model cycle

Further, it has been shown that the model's nonlinear objective function ensures a more realistic, more flexible and smoother response compared to linear programming models. Finally, the minimum data required for constructing the model are easily obtained. Overall, the SASEM model exhibits a range of characteristics that are generally desirable for agricultural policy models.

9.2. Contribution of the thesis

Traditional policy analysis has lacked an effective tool for dealing altogether with the hierarchical bi-level and multi-objective nature of the problems it tries to solve. By adapting and enhancing existing optimisation methods to develop modelling tools suitable for real-world policy planning tasks, this thesis makes a significant original contribution in the area of policy analysis and optimisation.

The hybridisation approach offers a significant advance in the methods available to solve hierarchical decision problems with two levels. It goes beyond sector modelling and "what if?" type of policy experiments to the next logical step namely, the challenge of attempting to find the best or near-best feasible policies. In addition, the ability of the model to handle multiple objective functions facilitates the investigation of trade-offs between non-commensurable and conflicting objectives typical in sustainability scenarios. The structure of the tradeoffs among objectives provides a quantitative means of evaluating the opportunity costs of alternative policy decisions in economic and non-economic terms. Thus, the model's output creates a basis for negotiation in situation where there are different points of view between various interest groups over proposed policy measures. A decision made in this manner is transparent, rational and quantitatively justified.

APOLO is the first model in the field of agricultural policy analysis which is capable of finding simultaneously *i*) Pareto-optimal or near Pareto-optimal solutions to multi-objective policy problems, *ii)* which production patterns result in these solutions and *iii*) the values of policy decision variables which lead to these production patterns.

There is a long list of potential applications of BLP models in the field of agricultural and natural resource economics. Price, tax rate, payment and quota specification is an essential part in setting up policy measures. For example, in the context of the water policy framework a BLP model such as APOLO would be particularly useful in determining "optimal" prices for water used in agriculture. Likewise, the model could "recommend" optimal payments for biofuel crops and/or a range of agri-environmental programmes. The approach would also be suitable for formulating BLP models investigating theoretical debates such as the issue of agricultural multifunctionality as a legitimate argument for policy intervention. For example, if non-market commodities comprise the model's objectives and market commodities its production activities, the existence and extent of jointness between the two could be assessed by examining the relationship between production activity levels and objective achievement levels for all solutions in the non-dominated set. The level of correlation found between a certain production activity and a non-marketable objective would provide a strong indication of whether producing the former results also in joint production of the latter or not.

9.3. Challenges for Improvements and Future Research

Validation of the APOLO model was done by testing it against simple but not trivial BLP problems. The fact that it had no difficulties in finding the global optimal solutions to these problems suggests that the method has potentials. In addition, the functionality of the method has been demonstrated through a real-world application. In all cases where the APOLO model was used it produced sensible and reliable results. Nevertheless, its performance should be further evaluated before using it with full confidence in real-world application.

It should be emphasised that the performance of the hybrid MOBLP algorithm is subject to the assumptions and limitations inherent in both the MOGA used for the leader problem and the model used for the follower. The choice of the “right” MOGA relates to its ability to find non-dominated fronts as close to the true Pareto-optimal front as possible. Moreover, these non-dominated fronts should contain diverse solutions evenly distributed along the front. The choice of the “right” follower model is to a large extent dependant of the type of the application aimed for. The remainder of the discussion takes these issues a bit further.

The literature suggests that MOGAs are able to solve different two-objective optimisation problems but they have yet to show their efficacy in handling adequately problems having more than two objectives (Deb et al., 2001). While the incorporation of many objectives may be desirable, at the same time it introduces two main disadvantages. With an increase in number of objective functions, the dimensionality of the objective space as well as the number of non-dominated solutions in the initial random population also increase. Both have serious implications in the initialisation stage, in fitness evaluation and consequently in the performance of the selection operators. The setting of some of the algorithm’s parameter values such as population size is also dependent on the number of objectives.

If all population members lie in the first non-dominated front, most MOEAs also assign the same (or similar) fitness to all solutions. When this happens, there is no selection advantage to any of these solutions. In the absence of any selection pressure for better solutions, recombination and mutation operators may not be very effective in finding better solutions. There are two ways to mitigate this problem, either use a large population size or modify the genetic algorithm. The first option offers only a temporary remedy since the computational burden of processing large number of solutions greatly slows the algorithm.

Algorithmic modification involves using a different evaluation criterion than assigning fitness based on the non-dominated rank of a solution. The amount of spread in objective space may be used to assign fitness. In this way, solutions that are closely packed in one part of the non-dominated front will not be favoured when compared to those that lie in less dense regions on the non-dominated front. However, this modification does not ensure that the search advances towards better solutions. This topic remains a challenge for future research in the field of multi-objective optimisation with evolutionary algorithms.

As mentioned above the capacity of MOBLP algorithm to handle policy problems is partly determined by the degree of realism captured by the model for the follower problem. For example, as already stated, in the SASEM model there is no sufficient technological, spatial and temporal disaggregation of production. Hence, not all the truly binding constraints are incorporated in the model's structure. For example, instead of monthly or seasonal constraints, an annual land constraint is used. Consequently, the model does not reflect the peak-season competition for resources which largely determine the cropping patterns. Because they work in similar spatial scales regional models are more suited to work in conjunction with biophysical models of agricultural systems. There exist sophisticated biophysical models that when integrated with economic agricultural models they can simulate the effect of land use, management and farming practices changes on chemical pollution, soil losses and biodiversity. The output of these models could provide more reliable metrics on objective achievement used by the ranking operator in the genetic algorithm.

Moreover, the SASEM model is essentially a profit maximisation based modelling approach. It assumes that farmers have a single (profit maximisation) rather than multiple objectives. Recent results from a DEFRA project commissioned to the School of Agriculture Policy and Development, University of Reading, (Garforth and Rehman, 2006) show that all farmers surveyed use a combination of different objectives to inform their decisions. Hence, to model the behaviour and motivations of farmers in responding to policy changes they apply a neat methodology developed by Yates (2006) which uses observations of past behaviour to estimate the relative importance of each objective to different farm types. This approach captures the behaviour of different farmer types and should be preferred in applications concerning the impact of policy change on distinct farmer groups.

The choice of the most suitable model for the follower’s problem should be application-specific. Depending on the scope of the application, the follower’s model can be a mathematical model of any kind (linear, PMP-based, farm, regional or sectoral, static or dynamic, single or multi-objective) if there is reason to believe that it better simulates the system under study. If macro effects are the focus of the application then partial or general equilibrium models may be used. Whatever the choice for the generic algorithm used for the leader’s problem or the modelling tool for the follower’s problem the MOBLP framework that this thesis proposes offers a robust approach for solving problems characterised by sequential execution of decisions between two independently interacting decision-making units.

Bibliography

Anandalingam, G., Mathieu, R., Pittard, L., & Sinha, R. 1989, "Artificial Intelligence Based Approaches for Hierarchical Optimization," in *Impact of Recent Computer Advances in Operations Research*, Sharda et al, ed., North-Holland, New York.

Anderson, K. 2000, "Agriculture's 'multifunctionality' and the WTO", *Australian Journal of Agricultural and Resource Economics*, vol. 44, no. 3, pp. 475-494.

Arfini, F. and Paris Q. 1995. A positive mathematical programming model for regional analysis of agricultural policies, *in* Sotte, E. (ed), *The Regional Dimension in Agricultural Economics and Policies*, Proceedings of the 40th EAAE Seminar, Ancona, Italy, 26-28, June 1995.

Arriaza, M. & Gomez-Limon, J. A. 2003, "Comparative performance of selected mathematical programming models", *Agricultural Systems*, vol. 77, no. 2, pp. 155-171.

Baldock, D. and Michell, K. 1995. Cross-compliance within the Common Agricultural Policy: A review of options for Landscape and nature conservation, Institute for European Environmental Policy, London, UK.

Bard, J. F. Solution Algorithms for the Government-Agriculture Bilevel Programming Model. Progress Report, prepared for Institut National de la Recherche Agronomique (INRA), Thivernal-Grignon, France by Department of Mechanical Engineering, University of Texas, Austin (1996).

Bard, J. F. 1998, *Practical Bilevel Optimization Algorithms and Applications* Kluwer Academic Publishers, Dordrecht, The Netherlands.

Barkaoui, A. & Butault, J. 2000, "Cereals and oilseeds supply within the EU, under AGENDA 2000: a Positive Mathematical Programming application", *Agricultural Economics Review*, vol. 1, no. 2, pp. 7-18.

Bauer, S. & Kasnakoglu, H. "Non-linear programming models for sector and policy analysis", *Economic Modelling*, vol. 7, pp. 275-290.

Bazarra, M.S., Jarvis, J.J. & Sherali, H.D. 1990. *Linear Programming and Network Flows* John Wiley & Sons, New York.

Bennett, D. A., Xiao, N. C., & Armstrong, M. P. 2004, "Exploring the geographic consequences of public policies using evolutionary algorithms", *Annals of the Association of American Geographers*, vol. 94, no. 4, pp. 827-847.

Burrell, A. 2001. Multifunctionality and agricultural trade liberalization, 77th Seminar of the European Association of Agricultural Economists, Helsinki, Finland, August 17-18, 2001.

Candler, W., Fortunyamat, J., & Mccarl, B. 1981, "The Potential Role of Multilevel Programming in Agricultural-Economics", *American Journal of Agricultural Economics*, vol. 63, no. 3, pp. 521-531.

Candler, W. & Norton, R. Multi-Level Programming and Development Policy. Working Paper No. 258, World Bank. Washington DC (May 1977).

Chadwick, L. 2005, *The Farm Management Handbook 2004/05*, 25 edn, Scottish Agricultural College, Edinburgh.

Chau, N. H. & de Gorter, H. 2005, "Disentangling the Consequences of Direct Payment Schemes in Agriculture on Fixed Costs, Exit Decisions, and Output", *American Journal of Agricultural Economics*, vol. 87, pp. 1174-1181.

Colson, B., Marcotte, P., & Savard, G. 2005, "A trust-region method for nonlinear bilevel programming: Algorithm and computational experience", *Computational Optimization and Applications*, vol. 30, no. 3, pp. 211-227.

Colson, B., Marcotte, P., & Savard, G. 2005, "A trust-region method for nonlinear bilevel programming: Algorithm and computational experience", *Computational Optimization and Applications*, vol. 30, no. 3, pp. 211-227.

Dantzig, G.B. *Linear Programming and Extension*. Princeton University Press, Princeton, New Jersey.

Day, R. H. 1963, "On aggregating linear programming models of production", *Journal of Farm Economics*, vol. 45, pp. 797-813.

Deb, K. 1999, "Multi-objective Genetic Algorithms: Problem Difficulties and Construction of Test Problems", *Evolutionary Computation*, vol. 7, no. 3, pp. 205-230.

Deb, K. 2001, *Multi-objective Optimisation using Evolutionary Algorithms* Wiley, West Sussex, England.

Deb, K., Agrawal, S., Pratap, A., & Meyarivan, T. 2002, "A Fast Elitist Non-Dominated Sorting Genetic Algorithm: NSGA-II", *IEEE Transaction on Evolutionary Computation*, vol. 6, no. 2, pp. 182-197.

Dempe, S. 2002, *Foundations of Bilevel Programming* Kluwer Academic Publishers, Dordrecht, The Netherlands.

Di, H. J. & Cameron, K. C. 2002, "Nitrate leaching in temperate agroecosystems: sources, factors and mitigating strategies", *Nutrient Cycling in Agroecosystems*, vol. 64, no. 3, pp. 237-256.

Dorsey, R.E. & Mayer, W.J. 1995. “Genetic Algorithms for Estimation Problems With Multiple Optima, Nondifferentialbility, and Other Irregular Features”, *Journal of Business & Economic Statistics*, Vol. 13, no. 1, pp. 53-66.

European Commission, 2000. Indicators for the integration of environmental concerns into the Common Agricultural Policy. Communication to the Council, the European Parliament, COM 2000 (20) Final, January 26, 2000. http://europa.eu.int/comm/agriculture/envir/com20/20_en.pdf.

Falconer, K. & Hodge, I. 2000, "Using economic incentives for pesticide usage reductions: responsiveness to input taxation and agricultural systems", *Agricultural Systems*, vol. 63, no. 3, pp. 175-194.

Flury, C., Gotsch, N., & Rieder, P. 2005, "Site-specific and regionally optimal direct payments for mountain agriculture", *Land Use Policy*, vol. 22, no. 3, pp. 207-214.

Forrest, S. 1993, "Genetic algorithms: principles of natural selection applied to computation", *Science*, vol. 261, pp. 872-878.

Fortuny-Amat, J., & McCarl, B. 1981. "A Representation and Economic Interpretation of a Two-Level Programming Problem", *The Journal of the Operational Research Society*, Vol. 32, no. 9, pp. 783-792.

Fraser, I. 2000, "An application of maximum entropy estimation: the demand for meat in the United Kingdom", *Applied Economics*, vol. 32, no. 1, pp. 45-59.

Garforth, C and Rehman, T. 2006. "Research to Understand and Model the Behaviour and Motivations of Farmers in Responding to Policy Changes (England)", Final Report, prepared for the Department for Environment, Food and Rural Affairs (DEFRA), by School of Agriculture, Policy and Development, University of Reading, Reading, UK.

Giampietro, M. 2004, *Multi-scale integrated analysis of agroecosystems* CRC Press, Boca Raton, Florida.

Golan, A., Judge, G., & Miller, D. 1996, *Maximum Entropy Econometrics Robust Estimation wiht Limited Data* Willey, West Sussex, England.

Goldberg, D. E. 1989, *Genetic Algorithms in Search, Optimization, and Machine Learning* Addison-Wesley, Reading, Massachusetts.

Goldberg, D. E., Deb, K., & Korb, B. "Don't worry, be messy", R. K. Belew & L. B. Booker, eds., Morgan Kaufmann Publishers, University of California, San Diego, pp. 24-30.

Harvey, D. R. 2003, "Agri-environmental relationships and multi-functionality: Further considerations", *World Economy*, vol. 26, no. 5, pp. 705-725.

Haubelt, C., Mostaghim, S., Teich, R., & Tyagi, A. 2003, "Solving hierarchical optimization problems using MOEAs", *Lecture notes in computer science*, vol. 2632, pp. 162-176.

Hazell, P. & Norton, R. 1986, *Mathematical programming for economic analysis in agriculture* Macmillan, New York.

Hazell, P. B. R. 1982, "Application of Risk Preference Estimates in Firm-Household and Agricultural Sector Models", *American Journal of Agricultural Economics*, vol. 64, no. 2, pp. 384-390.

Heckelei, T. 1997. Positive mathematical programming: Review of the standard approach, CAPRI-working paper 97-03, University of Bonn.

Heckelei, T. & Britz, W. 2000, "Positive Mathematical Programming with Multiple Data Points: A Cross-Sectional Estimation Procedure", *Cahiers d' économie et sociologie rurales*, vol. 57, no. 28, p. 50.

Heckelei, T., H.P. Witzke, and W. Henrichsmeyer (eds.). *Agricultural Sector Modelling and Policy Information Systems*, Proceedings of the 65th Seminar of the European Association of Agricultural Economists, Bonn, Germany, March 29-31, 2000. Wissenschaftsverlag VaukKiel, Kiel, Germany, 2001.

Heckelei, T. & Wolff, H. 2003, "Estimation of constrained optimisation models for agricultural supply analysis based on generilised maximum entropy", *European Review of Agricultural Economics*, vol. 30, no. 1, pp. 27-50.

Heckelei, T. 2005. Shadow Prices in PMP and consequences for Calibration and Estimation of Programming Models. Paper presented at the XIth International Congress of the European Association of Agricultural Economists. Copenhagen, Denmark August 24-27, 2005.

Helming, J.F.M., Peeters, L., and Veendendaal. P.J.J. 2001. Assessing the Consequances of Environmental Policy Scenarios in Flemish Agriculture. In Heckelei T., Witzke, H.P., and Henrichsmeyer, W. (eds.), *Agricultural Sector Modelling and Policy Information Systems*, Proceedings of the 65th Seminar of the European Association of Agricultural Economists, Bonn, Germany, March 29-31, 2000. Wissenschaftsverlag VaukKiel, Kiel, Germany, 2001.

Hennessy, T. 2000. "Modelling Farmer Response to Policy Reform: An Irish Example". In Heckelei T., Witzke, H.P., and Henrichsmeyer, W. (eds.), *Agricultural Sector Modelling and Policy Information Systems*, Proceedings of the 65th Seminar of the European Association of Agricultural Economists, Bonn, Germany, March 29-31, 2000. Wissenschaftsverlag VaukKiel, Kiel, Germany, 2001.

Henry de Frahan, B. 2005. PMP, Extensions and Alternative methods: Introductory Review of the State of the Art. Paper presented at the XIth International Congress of the European Association of Agricultural Economists. Copenhagen, Denmark August 24-27, 2005.

Hodge, I. 2000, "Agri-environmental Pelationships and the Choice of Policy Mechanism", *The World Economy*, vol. 23, no. 2, pp. 257-273.

Holland, J. 1975. *Adaptation in natural and artificial systems*, University of Michigan University Press, Ann Arbor, USA.

Horst, R. & Tuy, H. 1995, *Global Optimization: Deterministic Approaches*, Third edn, Springer-Verlag, Berlin.

Howitt, R. & Reynaud, A. 2003, "Spatial disaggregation of agricultural production data using maximum entropy", *European Review of Agricultural Economics*, vol. 30, no. 3, pp. 359-387.

Howitt, R. E. 1995, "Positive Mathematical Programming", *Americal Journal of Agricultural Economics*, vol. 77, pp. 329-342.

Howitt, R. E. 2005a. PMP Based Production Models - Development and Integration. Paper presented at the XIth International Congress of the European Association of Agricultural Economists. Copenhagen, Denmark August 24-27, 2005.

Howitt, R.E. 2005b. *Agricultural and Environmental Policy Models: Calibration, Estimation and Optimization*, Unpublished.

Hwang, C. L., Masud, A. S. N., Paidy, S. R., & Yoon, K. 1979, *Multiple Objective Decision Making - Methods and Applications: a State-of-the-Art Survey*. Springer-Verlag, Berlin.

Ignizio, J. P. 1978, "Review of Goal Programming - Tool for Multiobjective Analysis", *Journal of the Operational Research Society*, vol. 29, no. 11, pp. 1109-1119.

Jaynes, E. T. 1957a, "Information theory and statistical mechanics", *Physics Review*, vol. 106, pp. 620-630.

Jaynes, E. T. 1957b, "Information theory and statistical mechanics II", *Physics Review*, vol. 108, pp. 171-190.

Jones, G. 1982. *The effect of Milk Prices on the Dairy Herd and Milk Supply in the United Kingdom and Ireland.*, Institute of Agricultural Economics, University of Oxford, Oxford.

Jones, P. J., Rehman, T., Harvey, D. R., Marsh, R. B., Bunce, R. G. H., & Howard, D. C. 1995. *Developing LUAM (Land Use Allocation Model) and modelling CAP reforms.* CAS Paper 32. Centre for Agricultural Strategy, The University of Reading, Reading UK.

Judez, L., Chaya, C., Martinez, S., & Gonzalez, A. A. 2001, "Effects of the measures envisaged in "Agenda 2000" on arable crop producers and beef and veal producers: an application of Positive Mathematical Programming to representative farms of a Spanish region", *Agricultural Systems*, vol. 67, no. 2, pp. 121-138.

Judez, L., de Miguel, J. M., Mas, J., & Bru, R. 2002, "Modeling crop regional production using positive mathematical programming", *Mathematical and Computer Modelling*, vol. 35, no. 1-2, pp. 77-86.

Kim, C. & Kim, Y. 2000, "Economic Effects of Eco-taxation on Chemical Fertilizers", *Korean Journal of Agricultural Economics*, vol. 42, pp. 25-43.

Lancaster, K. 1966, "A new approach to consumer theory", *Journal of Political Economy*, vol. 74, pp. 132-157.

Lara, P. & Stancu-Minasian, I. 1999, "Fractional programming: a tool for the assessment of sustainability", *Agricultural Systems*, vol. 62, pp. 131-141.

Lee, S. M. & Clayton, E. R. 1972, "Goal Programming Model for Academic Resource Allocation", *Management Science Series B-Application*, vol. 18, no. 8, p. B395-B408.

Lence, S. H. & Miller, D. J. 1998, "Estimation of multi-output production functions with incomplete data: A generalised maximum entropy approach", *European Review of Agricultural Economics*, vol. 25, no. 2, pp. 188-209.

Leon, Y., Peeters, L., Quinqu, M., & Surry, Y. 1999, "The use of Maximum Entropy to estimate input-output coefficients from regional farm accounting data", *Journal of Agricultural Economics*, vol. 50, no. 3, pp. 425-439.

Liu, B. D. 1998, "Stackelberg-Nash equilibrium for multilevel programming with multiple followers using genetic algorithms", *Computers & Mathematics with Applications*, vol. 36, no. 7, pp. 79-89.

MAFF 2000. *Towards Sustainable Agriculture: A Pilot Set of Indicators*. London: MAFF Publications.

McCarl, B.A. & Apland, J. "Validation of linear programming models", *Southern Journal of Agricultural Economics*, vol. 68, pp. 155-164.

Marcotte, P. & Savard, G. 1991, "A Note on the Pareto Optimality of Solutions to the Linear Bilevel Programming Problem", *Computers & Operations Research*, vol. 18, no. 4, pp. 355-359.

Martin, R. P., Khalid, S., & Ahmed, S. March 2000.An Introduction to Linear Bi-Level Optimization. Working Paper No. 00-25. Lahore University of Management Sciences, Pakistan.

Matthews, K. B. 2001. *Applying Genetic Algorithms to Multi-objective Land-Use Planning,* PhD thesis, The Robert Gordon University.

Miettinen, K. M. 1999, *Nonlinear Multiobjective Optimization* Kluwer Academic Publishers, Boston.

Morgan, R. P. C. 1995, *Soil Erosion and Conservation,* Second edn, Longman Group, London.

Moxey, A. P., White, B., & O'Callaghan, J. R. 1995, "The Economic Component of NELUP", *Journal of Environmental Planning and Management*, vol. 38, pp. 21-33.

Munda, G. 2004, "Social multi-criteria evaluation: Methodological foundations and operational consequences", *European Journal of Operational Research*, vol. 158, no. 3, pp. 662-677.

Nilsson, H. 2004, "What are the possible influences affecting the future environmental agricultural policy in the European Union? An investigation into the main factors", *Journal of Cleaner Production*, vol. 12, no. 5, pp. 461-468.

Nix, J. 2004, *Farm Management Pockebook,* 32 edn, The Andersons Centre, Wye.

Norton, R. D. & Schieffer, G. W. 1980, "Agricultural sector programming models: A review", *European Review of Agricultural Economics*, vol. 7, no. 3, pp. 229-265.

O'Callaghan, J. R. 1996, *Land Use: the interaction of economics, ecology and hydrology* Chapman and Hall, London.

Oduguwa, V. & Roy, R. 2002, "Bi-level optimisation using genetic algorithm", *2002 IEEE International Conference on Artificial Intelligence Systems*, Proceedings pp. 322-327.

OECD 1997, *Environmental Indicators for Agriculture, vol. 1: Concepts and Framework.*, Publications Service, OECD, Paris.

OECD 1999, *Environmental Indicators for Agriculture, vol. 2: Issues and Design.*, Publications Service, OECD, Paris.

OECD 2001. *Multifunctionality: Applying the OECD Analytical Framework Guiding Policy Design*. OECD workshop in Paris, 2 – 3 July 2001.

OECD 2006, "Decoupling Illustrating some Open Questions on the Production Impact of Different Policy Instruments: No 416", *OECD Papers*, vol. 5, pp. 61-102.

Office of Technology Assessment 1995. Environmental Policy Tools: A User's Guide, Congress of the United States.

Oglethorpe, D. R. & Sanderson, R. A. 1999, "An Ecological-Economic Model for Agri-Environmental Policy Analysis", *Ecological Economics*, vol. 28, pp. 245-266.

Omar, B. & Blair, C. E. 1990, "Computational Difficulties of Bilevel Linear Programming", *Operations Research*, vol. 38, no. 3, pp. 556-560.

Onal, H., Darmawan, D. H., & Johnson, S. H. 1995, "A Multilevel Analysis of Agricultural Credit Distribution in East Java, Indonesia", *Computers & Operations Research*, vol. 22, no. 2, pp. 227-236.

Opschoor, J.B. and Vos, H.B. 1989. Economic Instruments for Environmental Protection, OECD.

Pacini, C., Wossink, A., Giesen, G., & Huirne, R. 2003, "Ecological-economic modelling to support multi-objective policy making: a farming systems approach implemented for Tuscany", *Agriculture, Ecosystems & Environment*, vol. 102, no. 3, pp. 349-364.

Paris, Q. 2001, "Symmetric Positive Equilibrium Problem: A Framework for Rationalizing Economic Behavior with Limited Information", *Americal Journal of Agricultural Economics*, vol. 83, no. 4, pp. 1049-1061.

Paris, Q. and Caputo, M.R., 2001. Sensitivity of the GME Estimates to Support Bounds, Working Paper No. 01-008, Department of Agricultural and Resource Economics, University of California, Davis.

Paris, Q. & Howitt, R. E. 1998, "An analysis of ill-posed production problems using maximum entropy", *American Journal of Agricultural Economics*, vol. 80, no. 1, pp. 124-138.

Piorr, H. P. 2003, "Environmental policy, agri-environmental indicators and landscape indicators", *Agriculture Ecosystems & Environment*, vol. 98, no. 1-3, pp. 17-33.

Prestegard, S.S. 2003. Policy Measures to Enhance a Multifunctional Agriculture: Applications to the WTO Negotiations on Agriculture. Paper presented at the International Conference “Agricultural policy reform and the WTO: where are we heading?”, Capri, Italy, June 23-26, 2003.

Rohm, O. & Dabbert, S. 2003, "Integrating agri-environmental programs into regional production models: An extension of positive mathematical programming", *American Journal of Agricultural Economics*, vol. 85, no. 1, pp. 254-265.

Romero, C. & Rehman, T. 2003, *Multiple Criteria Analysis for Agricultural Decisions*, 2nd edn, Elsevier, Amsterdam.

Romstad, E., Vatn, A., Rørstad, P.K. and Søyland, V. 2000. Multifunctional Agriculture. Implications for Policy Design. Report No 21, Agricultural University of Norway, Department of Economics and Social Sciences, Ås.

Rubenstein-Montano, B., Anandalingam, G., & Zandi, I. 2000, "A genetic algorithm approach to policy design for consequence minimization", *European Journal of Operational Research*, vol. 124, pp. 43-54.

Salvatici, L., Anania, G., Arfini, F., Conforti, P., De Muro, P., Londero, P., & Sckokai, P. 2001. "Recent developments in modelling the CAP: hype or hope?" In Heckelei T., Witzke, H.P., and Henrichsmeyer, W. (eds.), *Agricultural Sector Modelling and Policy Information Systems*, Proceedings of the 65th Seminar of the European Association of Agricultural Economists, Bonn, Germany, March 29-31, 2000. Wissenschaftsverlag VaukKiel, Kiel, Germany, 2001.

Shannon, C. E. 1948, "A mathematical theory of communication", *Bell System Technical Journal*, vol. 27, pp. 379-423.

Schaffer, J. 1985. Multiple objective optimisation with vector evaluated genetic algorithms. In Grefenstette (ed.), *Proceedings of the First International Conference on Genetic Algorithms*, Lawrence Earlbaum, Hillsdale, NJ, pp 93-100.

Soule, J. D. & Piper, J. K. 1992, *Farming in Nature's Image: An Ecological Approach to Agriculture.* Island Press, Washington.

Sumelius, A. 1997. Concerns related to possible effects of trade liberalization on landscape and biodiversity in the Nordic countries. *Proceedings of NJF-seminar no. 273. Agriculture, Trade and the Environment*. NJF-utredning/Rapport Nr. 117.

Swinbank, A. 2001. Multifunctionality: A European Euphemism for Protection? FWAG Conference, National Agricultural Centre, Stoneleigh, 29 November 2001.

Tiwari, D. N., Loof, R., & Paudyal, G. N. 1999, "Environmental-economic decision-making in lowland irrigated agriculture using multi-criteria analysis techniques", *Agricultural Systems*, vol. 60, pp. 99-112.

Topp, C. F. E. & Mitchell, M. M. 2003, "Forecasting the environmental and socio-economic consequences of changes in the Common Agricultural Policy", *Agricultural Systems*, vol. 76, pp. 227-252.

Umstatter, J. 1999, *Calibrating Regional Production Models Using Positive Mathematical Programming. An Agro-evnironmental Policy Analysis in Southwest Germany* Shaker Verlag, Germany.

Vatn, A. 1999. Integrated policy approaches to multifunctionality. Discussion paper D-18/2000, Department of Economics and Social Sciences, Agricultural University of Norway, Ås

Vicente, L. N. & Calamai, P. H. 1994, "Bilevel and Multilevel Programming - A Bibliography Review", *Journal of Global Optimization*, vol. 5, no. 3, pp. 291-306.

Wiborg, T., McCarl, B.A., Rasmussen, S. & Schneider, U.A. 2005. Aggregation and Calibration of Agricultural Sector Models through Crop Mix Restrictions and Marginal Profit Adjustments. Paper presented at the XIth International Congress of the European Association of Agricultural Economists. Copenhagen, Denmark August 24-27, 2005.

Wischmeier, W.C., and Smith, D.D. 1978. Predicting rainfall erosion losses - A guide to conservation planning. Agricultural Handbook 537, U.S. Gov. Print. Office, Washington, DC.

WTO 2001. *Ministerial Declaration*. Fourth WTO Ministerial Conference in Doha, Qatar from 9 November to 14 November 2001.

Yates, C. M. 2006, "A positive approach to estimating the weights for quadratic multiple objective programming", *Journal of Operations Research Society*, Advance online publication, 13 September 2006, pp. 1-9.

Yin, Y. F. 2000, "Genetic-algorithms-based approach for bilevel programming models", *Journal of Transportation Engineering-Asce*, vol. 126, no. 2, pp. 115-120.

Yin, Y. F. 2002, "Multiobjective bilevel optimization for transportation planning and management problems", *Journal of Advanced Transportation*, vol. 36, no. 1, pp. 93-105.

Yiridoe, E. K. & Weersink, A. 1997, "A Review and Evaluation of Agroecosystem Health Analysis: The Role of Economics", *Agricultural Systems*, vol. 55, no. 4.

Zadeh, L. A. 1963, "Optimality and Nonscalar-valued Performance Criteria", *IEEE Transaction on Automatic Control*, vol. 8, no. 1, pp. 59-60.

Zander, P. & Kächele, H. 1999, "Modelling multiple objectives of land use for sustainable development", *Agricultural Systems*, vol. 59, pp. 311-325.

Zeleny, M. 2005. "The Evolution of Optimality: De Novo Programming", *Lecture Notes in Computer Science*, vol. 3410, pp. 1-13.